

**PERCEPTIONS AND ATTRIBUTES OF ELECTRONIC THESES  
AND DISSERTATIONS (ETDS): A SOUTH AFRICAN STUDY.**

by

Neil Davies Evans

Dissertation submitted to the Faculty of Arts in fulfillment of the requirements for the degree of Master of Arts, in the Department of Communication Science, University of Zululand.

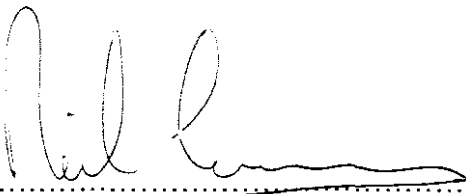
Supervisor : Professor G.M. Mersham

2006

## STATEMENT OF ORIGINALITY

This dissertation is an original piece of work by the author and has not been submitted elsewhere for a degree.

Signed:



.....

## **ACKNOWLEDGEMENTS**

I would like to offer a sincere thank you to my supervisor Prof. G.M. Mersham for his guidance throughout my study.

To my Dad who passed away last year this is dedicated to you. To family, friends and colleagues your encouragement and assistance also need to be acknowledged.

## TABLE OF CONTENTS

|                                                                         |           |
|-------------------------------------------------------------------------|-----------|
| STATEMENT OF ORIGINALITY .....                                          | 1         |
| ACKNOWLEDGEMENTS.....                                                   | 2         |
| TABLE OF CONTENTS .....                                                 | 3         |
| LIST OF TABLES .....                                                    | 5         |
| LIST OF FIGURES.....                                                    | 6         |
| <br><b>CHAPTER ONE: THE PROBLEM AND ITS SETTING .....</b>               | <b>7</b>  |
| 1.1 INTRODUCTION.....                                                   | 7         |
| 1.2 HISTORICAL OVERVIEW .....                                           | 8         |
| 1.3 PROBLEM STATEMENT .....                                             | 10        |
| 1.4 MOTIVATION FOR THIS STUDY .....                                     | 11        |
| 1.5 PURPOSE OF THIS STUDY .....                                         | 12        |
| 1.6 CONCLUSION.....                                                     | 12        |
| <br><b>CHAPTER TWO: LITERATURE REVIEW .....</b>                         | <b>13</b> |
| 2.1 PERCEIVED BENEFITS AND CONTROVERSIES<br>ASSOCIATED WITH ETDS .....  | 13        |
| 2.1.1 INTRODUCTION.....                                                 | 13        |
| 2.1.2 DUPLICATION OF RESEARCH .....                                     | 14        |
| 2.1.3 ACCESSIBILITY AND VISIBILITY .....                                | 14        |
| 2.1.4 IMPROVEMENT OF RESEARCH PUBLICATION<br>TIMEFRAMES.....            | 15        |
| 2.1.5 COST BENEFITS.....                                                | 15        |
| 2.1.6 ETD FORMATS AND CONTENT CODING .....                              | 16        |
| 2.1.7 PLAGIARISM.....                                                   | 17        |
| 2.1.8 TECHNOLOGY AND IMPLEMENTATION.....                                | 18        |
| 2.1.9 PRIOR PUBLICATION.....                                            | 18        |
| 2.2 THE COMMUNICATION MODEL AND ETD LIFE-CYCLE.<br>.....                | 19        |
| 2.2.1 INTRODUCTION.....                                                 | 19        |
| 2.2.2 THE BROADER CULTURAL, SOCIETAL AND<br>INTERNATIONAL CONTEXT ..... | 21        |
| 2.2.3 THE COMMUNICATOR.....                                             | 23        |
| 2.2.4 ENCODING.....                                                     | 24        |
| 2.2.5 THE MESSAGE, SIGNS AND SYMBOLS .....                              | 25        |
| 2.2.6 THE MEDIUM .....                                                  | 25        |
| 2.2.7 THE RECIPIENT(S).....                                             | 26        |
| 2.3 ASPECTS OF INFORMATION MODELING .....                               | 27        |
| 2.3.1 INTRODUCTION.....                                                 | 27        |
| 2.3.2 WHAT IS AN INFORMATION MODEL .....                                | 28        |
| 2.3.3 THE THREE-TIERED STRUCTURE OF AN<br>INFORMATION MODEL .....       | 28        |
| 2.3.4 STATIC AND DYNAMIC INFORMATION MODELS .....                       | 29        |
| 2.3.5 ORGANIZATIONAL STRUCTURE .....                                    | 32        |
| 2.3.5.1 Hierarchical .....                                              | 32        |
| 2.3.5.2 Nonlinear .....                                                 | 33        |

|                                                                         |                                                                             |           |
|-------------------------------------------------------------------------|-----------------------------------------------------------------------------|-----------|
| 2.3.5.3                                                                 | Linear .....                                                                | 34        |
| 2.3.5.4                                                                 | Database-Driven .....                                                       | 35        |
| 2.3.6                                                                   | LIBRARIES AND ARCHIVING ETDs .....                                          | 35        |
| 2.3.7                                                                   | METADATA .....                                                              | 37        |
| 2.3.8                                                                   | CATALOGUING .....                                                           | 38        |
| 2.3.9                                                                   | INFORMATION RETRIEVAL .....                                                 | 39        |
| 2.3.10                                                                  | FULL TEXT SEARCHING .....                                                   | 39        |
| 2.3.11                                                                  | THE EXTENSIBLE MARKUP LANGUAGE (XML) .....                                  | 40        |
| 2.3.12                                                                  | MULTIMEDIA .....                                                            | 42        |
| 2.3.13                                                                  | SOFTWARE OPTIONS FOR HOSTING ETDs AND<br>PROVIDING THE USER INTERFACE ..... | 43        |
| 2.4                                                                     | SUMMARY .....                                                               | 44        |
| <b>CHAPTER THREE: METHODOLOGY .....</b>                                 |                                                                             | <b>45</b> |
| 3.1                                                                     | INTRODUCTION .....                                                          | 45        |
| 3.2                                                                     | OBJECTIVES .....                                                            | 46        |
| 3.3                                                                     | THE SURVEY INSTRUMENT .....                                                 | 46        |
| 3.4                                                                     | PROCEDURE .....                                                             | 47        |
| 3.5                                                                     | SAMPLE .....                                                                | 47        |
| 3.6                                                                     | CONCLUSION .....                                                            | 48        |
| <b>CHAPTER FOUR: RESULTS AND DATA INTERPRETATION .....</b>              |                                                                             | <b>49</b> |
| 4.1                                                                     | INTRODUCTION .....                                                          | 49        |
| 4.2                                                                     | RESULTS AND DATA INTERPRETATION .....                                       | 49        |
| 4.2.1                                                                   | SAMPLE .....                                                                | 49        |
| 4.2.2                                                                   | DEMOGRAPHICS .....                                                          | 51        |
| 4.2.3                                                                   | RESEARCH SOURCES .....                                                      | 53        |
| 4.2.4                                                                   | PERCEPTIONS OF ETDs .....                                                   | 56        |
| 4.2.5                                                                   | DUPLICATION OF RESEARCH .....                                               | 58        |
| 4.2.6                                                                   | ACCESSIBILITY AND VISIBILITY .....                                          | 60        |
| 4.2.7                                                                   | IMPROVEMENT OF RESEARCH PUBLICATION<br>TIMEFRAMES .....                     | 63        |
| 4.2.8                                                                   | COST BENEFITS .....                                                         | 64        |
| 4.2.9                                                                   | ETD FORMATS AND CONTENT CODING .....                                        | 66        |
| 4.2.10                                                                  | PLAGIARISM .....                                                            | 69        |
| 4.2.11                                                                  | TECHNOLOGY AND IMPLEMENTATION .....                                         | 69        |
| 4.2.12                                                                  | PRIOR PUBLICATION .....                                                     | 70        |
| 4.2.13                                                                  | SOFTWARE APPLICATIONS USED .....                                            | 70        |
| 4.3                                                                     | CONCLUSION .....                                                            | 71        |
| <b>CHAPTER FIVE: SUMMARY, CONCLUSIONS AND<br/>RECOMMENDATIONS .....</b> |                                                                             | <b>72</b> |
| 5.1                                                                     | INTRODUCTION .....                                                          | 72        |
| 5.2                                                                     | SUMMARY .....                                                               | 73        |
| 5.2.1                                                                   | Sample .....                                                                | 73        |
| 5.2.2                                                                   | Demographics .....                                                          | 73        |
| 5.2.3                                                                   | Research sources .....                                                      | 74        |
| 5.2.4                                                                   | Perceptions of ETDs .....                                                   | 74        |
| 5.2.5                                                                   | Software applications used in the preparation of TDs .....                  | 75        |
| 5.3                                                                     | CONCLUSIONS .....                                                           | 75        |

|              |                      |    |
|--------------|----------------------|----|
| 5.4          | RECOMMENDATIONS..... | 77 |
| BIBLIOGRAPHY | .....                | 81 |
| APPENDICES   | .....                | 86 |

## LIST OF TABLES

|                                                                                                                                                                                                  |    |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Table 2.1 in appendix 1 shows the recommended formats from Virginia Tech website (2006).                                                                                                         | 86 |
| Table 4.1 Tertiary Institutions where participants of the electronic survey worked or studied.                                                                                                   | 50 |
| Table 4.2 Occupation Frequencies.                                                                                                                                                                | 51 |
| Table 4.3 Gender of participants.                                                                                                                                                                | 52 |
| Table 4.4 Results to Q8. Have you used online databases to obtain research material?                                                                                                             | 53 |
| Table 4.5 Results to Q9. If you have, what did you access/download?                                                                                                                              | 53 |
| Table 4.6 shows the response to Q11: Do you find online databases more efficient to obtain research literature than paperbound journals?                                                         | 54 |
| Table 4.7 Sources of research material.                                                                                                                                                          | 55 |
| Table 4.8 shows the response to Q12: Are you familiar with electronic theses and dissertations (ETDs)?                                                                                           | 56 |
| Table 4.9 shows the response to Q16: If your institution does not accept ETDs, do you think it should?                                                                                           | 56 |
| Table 4.10 shows the response to Q25: Do you think by adding multi-media content (graphics, video, sound etc.) to theses and dissertations it will improve the expression of Graduate Education? | 57 |
| Table 4.11 shows the response to Q27: Would you as an author have any reservations on adding your thesis or dissertation to an online database which everybody can access?                       | 57 |
| Table 4.12 Concerns about publishing a thesis or dissertation (TD) on the Internet for everybody to access.                                                                                      | 58 |
| Table 4.13 Results of Q30 of the electronic survey.                                                                                                                                              | 59 |
| Table 4.14 Results of Q26 of the electronic survey.                                                                                                                                              | 61 |

Table 4.15 shows the response to Q31: Do you prefer to print downloaded literature to read? 71

## LIST OF FIGURES

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| Figure 2.1 Mersham's Graphic Communication model for development.         | 19 |
| Figure 2.2 Proposed life-cycle of an ETD.                                 | 20 |
| Figure 2.3 Hackos's Information Model from the Gilbane report (2002:4).   | 29 |
| Figure 2.4 Hierarchical organisational structures in Windows Explorer.    | 30 |
| Figure 2.5 Dynamic search functions within Windows Explorer.              | 31 |
| Figure 2.6 Broad and shallow Hierarchical system from Coorough (2001:40). | 33 |
| Figure 2.7 Narrow and deep Hierarchical system from Coorough (2001:41).   | 33 |
| Figure 2.8 Nonlinear organizational structures from Coorough (2001:41).   | 34 |
| Figure 2.9 Linear organizational structures from Coorough (2001:41).      | 35 |
| Figure 4.1 Tertiary education levels of participants.                     | 51 |
| Figure 4.2 The race of the participants.                                  | 52 |
| Figure 4.3 Results of Q23.                                                | 55 |
| Figure 4.4 Results of Q29.                                                | 58 |
| Figure 4.5 Illustrating the results of Q29 of the electronic survey.      | 64 |

## **CHAPTER ONE: THE PROBLEM AND ITS SETTING**

### **1.1 INTRODUCTION**

This thesis will focus on the social perceptions toward electronic theses and dissertations (ETDs) in South Africa and the attributes of these post-graduate studies that characterise access to information in the information age.

According to Figueres (2003:xi), the information revolution has globally impacted on the way institutions and organizations do business and the way those with access to information and communication technologies (ICTs) go about their daily lives. He goes on to acknowledge that ICTs offer a remarkable opportunity and set of tools for achieving progress in the broad spectrum of development challenges facing Africa including the promotion of its intellectual resources, which Ubogu (2003:70) states has been a problem for many years.

The website titled: Africa on the Internet. Starting Points for Policy Information (2004) expands on this by stating that for tertiary institutions, organizations and countries, failure to make the Internet connection in all aspects of communication will be a certain recipe for increasing marginalization within the 21st century. While the Global e-Schools community initiative website (2004) states that in a learning environment spreading ICTs, required for digital information processing has the capacity to improve the productivity of teachers and administrators, extend education and research to hard-to-reach places and provide students with access to content and tools for learning - more efficiently and at lower cost than traditional methods (Global e-Schools community: 2004).

For academic researchers an essential aspect is the dissemination of the findings arising from their research. This communication of information will hopefully benefit all for which the research was primarily intended while giving intellectual recognition to the researcher through the publication of the

findings. The traditional dissemination method of academic research has been primarily through publication of peer-reviewed manuscripts in journals. Unfortunately, according to Schweig *et al* (2001) this process often delays the release of important information for several months, often years and access to the full article is restricted to subscribers of the journal thereby reaching a relatively limited audience. Recognizing this, many scientific organizations *send out newsletters or use other mechanisms to fast-track important news events or time-critical results* (Schweig *et al*: 2001).

Lawrence (2001) states that improvements to the accessibility of scientific literature, allowing scientists to locate more relevant research within a given time, have the potential to dramatically improve communication and progress in science. Through the World Wide Web, scientists now have very convenient access to an increasing amount of literature that previously required trips to the library, inter-library loan delays, or substantial effort in locating the source (Lawrence: 2001).

For most Masters and Doctorate students in South Africa their theses and dissertations (TDs) will be printed a number of times then hard-bound for archiving in the institutions library, presenting to the supervisor and for personal use. This traditional method of TD dissemination reaches a fairly limited and normally local audience. However the emergence of electronic publishing, the worldwide trend toward accessing and sharing digital content through the Internet and increasing numbers of technologically literate people has led to the greater electronic dissemination of postgraduate academic TDs. Hosted in digital libraries, these electronic theses and dissertations (ETDs) *provide an archive of knowledge, which can easily be accessed by the local and international research community and diffused, across all fields of research.*

## **1.2 HISTORICAL OVERVIEW**

Fox (2001:38-39) explains how the initial concept of ETDs was first discussed at a meeting in the United States in Michigan in 1987. Arranged by the University of Microfilms (later known as UMI), and attended by

representatives of Virginia Tech, University of Michigan, SoftQuad, and ArborText. Southeastern Universities Research Association (SURA) funded the first pilot project in 1996, which was led by a team at Virginia Tech. He adds that as the idea of ETDs spread to other countries the Networked Digital Library of Theses and Dissertations (NDLTD) an international non-profit educational organization was established to co-ordinate and network collections of ETDs. According to Fox (2001:38-39) since its first annual meeting in 1998 the interest in ETDs has grown considerably. Currently there are two hundred and fifty six (256) institutional members with four (4) South African members being Rhodes University, University of Witwatersrand, University of the Free State and University of Pretoria (NDLTD website: 2005)

In Africa the Association of African Universities (AAU) implemented the database of African theses and dissertations (DATAD) pilot project from February 2000 to September 2003 (DATAD; 2003:4).

According to the final report the project aimed at exposing the research output of higher education institutions in Africa to the wider research community. The report further states that graduate TDs completed in African institutions are rarely indexed in major databases, nor do they feature much in the international literature. As a result the research they carry out has received little recognition (DATAD; 2003:5). The report states that a very different scenario exists in developed countries where graduate TDs are well preserved and indexed, and scholars and students worldwide have relatively easy access to them in print or digital form (DATAD; 2003:5). The AAU recognized that the digital communication revolution could provide a worldwide accessible, well-documented source of information on African TDs, through the DATAD project, its members have begun to come to terms with this challenge (DATAD; 2003:5).

According to its website, Rhodes University investigated the holding of electronic copies of theses and dissertations during 1996. The website stated that after several options were considered, the Senate, in April 1997, approved the option of requesting both print and digital copies of theses from

students. Rhodes was invited to join the NDLTD initiative, which the University duly did in May 1997 (Rhodes University: 2005).

According to the University of Pretoria's website, its UPeTD initiative was launched in July 2000 with the objective to create the necessary infrastructure and resources to allow postgraduate students to publish their theses or dissertations on the Internet in a well managed environment which will make it accessible to the international research community. The website also stated that UPeTD support open access and also forms part of the NDLTD (University of Pretoria: 2005).

At the University of the Witwatersrand a national ETD workshop was held on 16<sup>th</sup> and 17<sup>th</sup> of October, 2003 and examined various issues that should result in successful implementation of institutional as well as national ETD programmes (ETD Pilot Project, Final Report: 2004). According to their website the general rules of the University now requires every postgraduate student to submit two final, corrected copies of the dissertation or research report or thesis in a printed format as well as a final, corrected copy in electronic format. The website further states that the availability of Wits theses and dissertations electronically will undoubtedly lead to increased visibility of the institution internationally (University of Witwatersrand: 2005).

### **1.3 PROBLEM STATEMENT**

The wealth of knowledge and experience contained within TDs of academics was generally available in paperbound formats or as microfilm or microfiche. Obtaining these formats is normally a lengthy and sometimes expensive process, leading to restrictive access especially in institutions with limited research budgets.

In Mersham's (1995:55) graphic communication model, a vital part of the process is when the communicator manifests his or her ideas into the central position in the model, the message. In order for the communicator to transfer thoughts, information, feelings and attitudes, Mersham argues that we need to

manifest them in some concrete form as explicit information; otherwise they remain 'inner thoughts' or tacit information and not shared communication, perceptible to the senses through some or other medium. He also argues, in the footsteps of McLuhan, that the medium shapes the message. Thus the centrality of the message and its form (the medium in which it is encoded) are important in considerations in the communication or dissemination of research.

According to Fox (2001:21) universities should have curricula in place to ensure that students can create electronic documents that convey their knowledge and understanding, and demonstrate their ability to participate in the scholarly communication process. One assumes that postgraduate scholars have the required language and technical skills to produce ETDs and are doing so already using word processor applications in order to submit their TDs in the required paperbound format. However, the general lack of electronic publishing, indexing, and hosting of ETDs in most South African Universities denies scholars and the country the full benefits associated with accessing local research documentation through ICT networks.

Dalton (1999:3) noted that certain issues or 'controversies' had been identified in the discussions surrounding electronic publishing in general, and electronic dissertations specifically these will be addressed in chapter two. In addition, in developing countries such as South Africa, the question of access arises. These issues will be investigated and described within a South African perspective in this study.

#### **1.4 MOTIVATION FOR THIS STUDY**

According to Kling and McKim (1999:2) electronic publishing opportunities exist today in a variety of electronic journals and Web based collections, and scholarly publishing practices, especially those related to electronic publishing, are rapidly evolving. They go on to characterize three dimensions of scholarly publishing as a communicative practice, namely publicity, access, and trustworthiness. For the latter they suggested peer-reviewing provide a

valuable function in scholarly communication that can not effectively be replaced by self-published articles in electronic media (Kling and McKim; 1999:2).

According to Fox (2001:14) the preparation of ETDs allows the new generation of scholars to present their research results using the most appropriate format, composition, and content - all of which will continuously be re-defined as information and communication technologies (ICTs) and student knowledge advance (Fox; 2001:14).

The Robert Gordon University website (2006) acknowledge that during the process of creating ETDs postgraduate students are likely to improve their IT skills while expressing their research results in innovative and flexible ways that would not be possible if they were limited to paper based reports (see <http://www.lis.uzulu.ac.za/etd/etdcommodel.html> for an example of a proposed ETD communication model for this study).

## **1.5 PURPOSE OF THIS STUDY**

Given the lack of research on the perceptions and attributes of ETDs among Institutions of Higher Education in South Africa and the growing impact of ICTs on academic work, this research will contribute to a body of knowledge in this area, and hopefully encourage greater use of ETDs in the future.

## **1.6 CONCLUSION**

I conclude for this chapter that for TDs to be fully utilised in the information age changes in the traditional way they are produced, disseminated and accessed need to be recognized by tertiary institutions. The leading institutions highlighted in the historical overview will hopefully set the benchmark for others to follow.

## **CHAPTER TWO: LITERATURE REVIEW**

### **2.1 PERCEIVED BENEFITS AND CONTROVERSIES ASSOCIATED WITH ETDS**

#### **2.1.1 INTRODUCTION**

According to Ubogu (2003:1) the number of ETDs in Africa is increasing slowly with the majority of initiatives coming from South Africa. Ubogu (2001:33,249-259) is surprised that some institutional leaders in South Africa have not realized that the digital revolution makes it necessary to use current digital and publishing technologies not only in the provision of information but also in the preservation and dissemination of intellectual TDs.

Fox (2001:30) states that many universities have little or no involvement in digital library efforts, yet most will be involved in some project as part of, or in addition to, conventional library activities. He goes on to highlight the value of digital libraries particularly in distance education, and points out that by establishing an ETD program, a university automatically evolves into the digital library era.

Four perceived benefits of ETDs were identified in the literature reviewed and they will be discussed after the introduction of this chapter, namely:

- Minimize duplication
- Improve accessibility and visibility
- Improvement of research publication timeframes
- Cost benefits

Dalton (2001:3) identifies the following issues or controversies surrounding electronic publishing in general and ETDs in particular:

- the question of long-term preservation and transfer of content to future formats;
- concerns about how electronic access might facilitate plagiarism;
- the need for training graduate students to use the technology;

- implementation strategies needed to launch a project at academic institutions.

Dalton (2001:3) goes on to state that one of the most important issues identified was the level of awareness and acceptance of electronically distributed TDs by the scholarly publishing community and the question of prior publication, where derivative articles or portions of the research was being submitted for publication to journals. All these controversies will also be discussed in this chapter after the benefits.

### **2.1.2 DUPLICATION OF RESEARCH**

One of the arguments of advocates of ETDs is that they aid the process of searching for completed research and therefore prevent the duplication of unnecessary research efforts. Garfield (1994) stresses that researchers should get into the habit of literature searching to avoid the accidental duplication of research and the wasted time, effort, and funds this involves.

Fox (2001:16) states that one of the benefits of ETDs is that they can help minimize duplication of research as students are sometimes unaware of the findings of others mainly because there is not a complete database of research conducted that can be accessed.

### **2.1.3 ACCESSIBILITY AND VISIBILITY**

Fox (2001:32) states that access to ETDs may be the only option to those in developing countries who cannot afford to make purchases of TDs, who cannot wait for expensive delivery of copies through inter library loan, who cannot attend the multitude of conferences that demand the considerable travel expenses, or who cannot subscribe to expensive journals that may only have short summaries of thesis or dissertation results.

Sawyerr in Ubugo (2001:250) states that Africa needs to improve the accessibility and visibility of its intellectual TDs when he says:

“The absence of systematic compilation and indexing of postgraduate theses and dissertations in most African universities often means that

the countries for which such research is primarily undertaken are not able to benefit fully from it. But the loss goes beyond this. To the extent that literature searches have come to be conducted principally on the Internet, the absence of African theses and dissertations (and other intellectual production) from the major international databases reduces potential contribution of African research to worldwide knowledge while denying African scholarship the recognition it deserves”.

#### **2.1.4 IMPROVEMENT OF RESEARCH PUBLICATION TIMEFRAMES**

According to Sheth *et al* (1999) speed and distribution will characterize every aspect of most business and organizational undertakings in this millennium. Organizations will be challenged to deliver products and services at an ever-increasing pace in the worldwide marketplace. The question arises as to whether institutions of higher learning will or should respond to this need for faster research and distribution of research findings by moving towards ETDs.

Fox (2001:18) mentions that ETDs can be managed through automated procedures taking advantage of modern networked information systems. He states that checking of submissions and cataloging is sped up, moving and handling of paper copies is eliminated, and delays in binding are removed. He also states that the time between submission and graduation can be reduced, and ETDs can be made available for access within days or weeks rather than months.

#### **2.1.5 COST BENEFITS**

Kling and McKim (1999:11) state that posting a document in an unrestricted site on the Web potentially expands its readership to millions of people for little or no marginal cost.

Fox (2001:16) predicts that with a worldwide accessible collection of ETDs, students can quickly search for works related to their interest from anywhere in the world, and in most cases access and benefit from those studies without incurring any cost. Fox (2001:19) continues to spell out the cost benefits when

he says that ETD submission over networks has zero cost, which compares favourably with the high charges otherwise required to print, copy, or publish TDs using paper or other media forms.

### **2.1.6 ETD FORMATS AND CONTENT CODING**

Matthews and Wiggins (2001:41) state:

*"Standards must address more than markup and metadata: they must address format issues. Otherwise, a future scholar might retrieve a thesis via virtual interlibrary loan only to find that none of the content can be deciphered due to format obsolescence."*

Mersham's (1995:55) graphic communication model employed in this study draws attention to the process of coding and decoding the content or message in the communication process. Coding can be differentiated at two levels. The first takes place at the level of linguistic codes or language (see Mersham 1999:18-19) such as English, Afrikaans or Zulu. The second level has to do with technical encoding. This brings into focus the question of software and formats that are used to 'encode' the 'message' (research). The production of rich multimedia ETDs might require knowledge of technical codes and encoding (e.g. HTML, XML, internal targeting, hyper-linking, etc). Mersham (1999:17) refers to "restricted codes" as codes that are culturally or technically bound in their usage. Simply put, if one has access to code, one is 'in the know' and 'empowered', and conversely if one does not know the code or how to encode or decode, then one remains 'locked out' of the 'message'.

The University of Kentucky's (2006), in its ETD pilot programme, draws attention to format problems in the Frequently Asked Questions section of its website:

**Which file formats are accepted for the ETDs?** "During the pilot program, we are asking that you submit the text portion of your document in PDF. Acceptable formats for multimedia "enhancement" files (pictures, sound...) are listed in the Technical Guidelines section. We are committed to migrating the text portions of all ETDs to maintain

accessibility in the future. Multimedia files may not successfully migrate.”

**Why is PDF required?** “There is currently no single preferred electronic format for ETD submission. Most institutions with ETD programs have selected PDF because it is easy for the student to create, easy for the institution to distribute, easy for readers to access (a PDF viewer is available free), and it mimics a traditional paper document when it is printed (consistent page numbering, it looks like a traditional paper document, etc.) PDF is not a very good archival format, however. The UK ETD committee is exploring other formats. “

### **2.1.7 PLAGIARISM**

Guédon (2001:93) states that the many people tend to think that because a digitized TD is easily copied in part or in whole, it can be easily plagiarized, and therefore that it is better to keep TDs offline.

Guédon (2001:93) however, argues to the contrary:

“...the invention of the Philosophical Transactions (1665) by Henry Oldenburg, the Secretary to the Royal Society in London, was motivated by the issue of intellectual property. Oldenburg reasoned that if the research results of Scientist X. were printed in a journal (after being certified as being of good quality and original) and that journal was made widely available through the multiplication of copies, then Scientist. X would have a better chance to lay ownership claims than if he/she held back these results. By apparently giving away the results of his/her work, a scientist ensures his/her intellectual property most effectively. The ability to compare new results to already published work makes plagiarism a very risky business at best.”

Guédon (2001:94) proposes that ETDs available online will be more protected than TDs that remain poorly catalogued and unavailable outside the institution from which they are issued. He advises that visibility offers a very efficient

way to protect intellectual property and prevent plagiarism rediscovering Oldenburg's wisdom when it comes to scientific intellectual property.

### **2.1.8 TECHNOLOGY AND IMPLEMENTATION**

Figueres (2003: xi) states that the United Nations ICT Task Force identified Africa as the most unconnected in an increasingly connected world.

According to the Wits ETD Pilot Project's final report (2004:1) the concept of digital libraries is no longer new however many institutions were said to lack the financial resources necessary to implement an ETD project. The report (2004:2) mentioned much enthusiasm for the archiving and dissemination of ETDs, calling it a solution that would work well in the challenging environment of African universities, which have always had fewer resources than their non-African counterparts. The report (2004:2) foresees that with the support from institutions, national bodies and international agencies, digital libraries and ETDs are poised to make a significant impact on tertiary education in Africa.

### **2.1.9 PRIOR PUBLICATION**

Dalton (2001:3) notes that one of the most important issues identified at the second Symposium on Electronic Theses and Dissertations in May of 1999 in Virginia was the level of awareness and acceptance of electronically distributed TDs by the scholarly publishing community and the question of prior publication, where derivative articles or portions of the research was being submitted for publication to journals.

The following relevant question regarding prior publication was obtained from the University of Kentucky's UK-ETD website containing frequently asked questions: "Should I submit an ETD if I want to publish my thesis later?" The website states that this question is often cited as a reason students do not want to submit an ETD. The website goes on to quote McMillan (2001) who found that eighty six percent (86%) of publishers surveyed did not consider an ETD as being a 'prior publication'.

According to the Virginia Tech website, journals sometimes insist strictly on first publication and might well interpret WWW availability as synonymous to prior publication. They recommend identifying in advance the journals to which one might wish to submit ones article/s and discover their position on this question.

Dalton (2001:3-4) states that in response to these concerns, Virginia Tech offered students several options for restricting access to their work. Students could chose to offer full access to their ETD without any restrictions; restrict access to the campus community; restrict access for a specific period of time; or to restrict access entirely, which was desirable in specific cases such as with patent applications pending.

## 2.2 THE COMMUNICATION MODEL AND ETD LIFE-CYCLE

### 2.2.1 INTRODUCTION

To provide a theoretical foundation for the study, a development communication model is used heuristically to discuss the life-cycle of an ETD.

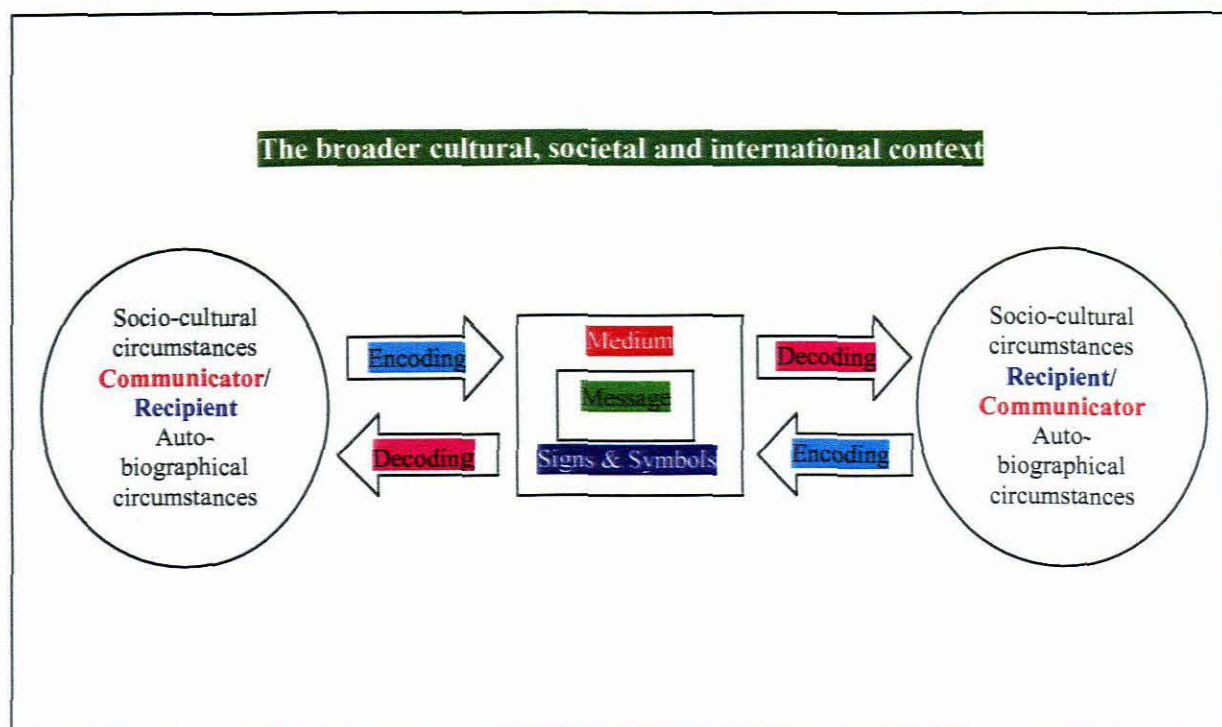


Figure 2.1 Merisham's Graphic Communication model for development.

The elements of Mersham's (1995:55) graphic communication model for development employed for this purpose are:

- The broader cultural, social and international context.
- The communicator
- Encoding
- Message, signs, symbols
- The medium
- Decoding
- The recipient(s)

Harper (2004) proposes the following Lifecycle of an ETD:

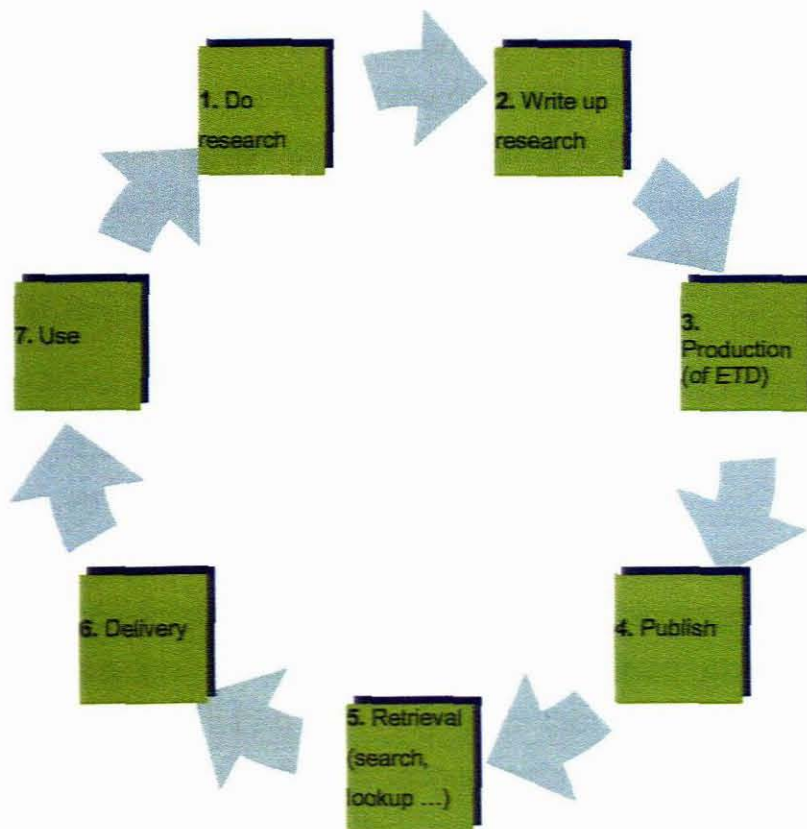


Figure 2.2 Proposed life-cycle of an ETD.

If we use Mersham's (1995:55) communication model and Harper's (2004) ETD life-cycle then we can explain the different aspects of communication in a methodical way with regard to the life-cycle of ETDs. Later in the chapter we

will look at an information model in the same light and steps three to six in more detail.

In step one of figure 2.2 the researcher or communicator conducts research using various information sources including online databases and the Web. In step two the communicator uses the prescribed software applications to write up, analyse and present his / her research. In Mersham's (1995:55) communication model this is represented as encoding the message. Step 3 can be seen as the conversion to a PDF or other acceptable formats before the ETD is catalogued and published online seen as step four in figure 2.2. The online environment represents the medium of the message which for ETDs is the Internet. The recipient(s) within Mersham's (1995:55) communication model which represent fellow academics, researchers etc. now search or lookup ETDs through a web user interface in step five of figure 2.2, and the delivery of the message or ETD constitutes the web server hosting the ETDs providing the download. The recipient(s) will then use the research by decoding the message using the appropriate software applications and language skills.

### **2.2.2 THE BROADER CULTURAL, SOCIETAL AND INTERNATIONAL CONTEXT**

Mersham's (1995:55) communication model indicates the importance of analysing communication interactions between communicators and recipients in terms of their broader cultural, societal and international contexts. In figure 2.1 the box (context) that surrounds the triptych of the communicator, message and the recipient represents the broader cultural, societal and international context that communicators to varying degrees find themselves in.

When publishing research on the Internet, academics expose their work to a much broader audience which Mersham (1995:57) refers to as the international context. He argues that taking this into account when encoding the message is "central to the passage or non-passage of innumerable

messages". The example can be given of different languages constituting different codes that recipients can decode or not. He refers to the 'global conversation' that characterises modern ICT activities. Through the dissemination of ETDs, scholarly communication puts 'global conversation' into practice when messages are archived on an "interoperable" database such as the NDLTD which has a large international audience.

The experience and circumstances of communicators in this international context, their world view, and the powerful social trends such as the '*internationalisation of communication*', have influential effects on the communication of research ideas within the international context.

Through live television, radio, telephones and the Internet we experience instantaneous communication over extended distances (Bowers: 2003). One can argue that this 'instantaneousness', a characteristic of contemporary 'global culture', should also apply to accessing ETDs over the Internet. Other, more traditional ways of accessing this type of research is far from instantaneous - for example the delivery of a TD via inter-library loan can take many weeks.

The different levels of access, skills and resources of the recipients of research information obviously means that different methods and media of communication will continue to be required for some time and paper based TDs will still be required. However, the greater demand for instantaneous communication methods to obtain information will impact on how and when academic research can be accessed. Academic institutions should recognize these preferred methods of communication by providing instantaneous access to their postgraduate knowledge while pursuing some form of credit for themselves and the author.

When disseminating research a communicator manifests his message in some format that can be accessed and understood by the recipient/s or intended audience. For this to be effective communication both parties need

to acknowledge the others socio-cultural and auto-biographical circumstances.

At the cultural level, South Africa's cultural groupings and their languages will need to be taken into account when communicating research findings to recipients with different languages. Linguistic software applications might in near future provide a valuable way to translate ETDs into several of the eleven (11) home languages recognized in South Africa. At the social level, issues surrounding disparities in education, development and access come to the fore when restructuring academic programmes to empower previously disadvantaged learners with sufficient ICT skills to participate in the ETD life-cycle. At the societal level South African national development objectives, and the role of ICTs should play in these objectives need to be highlighted and implemented. At the international level, issues concerning the 'digital divide' between countries, and how this affects their competitive edge in the worldwide research market need to be acknowledged by tertiary institutions seeking international collaboration and sponsorship.

### **2.2.3 THE COMMUNICATOR**

This aspect of the model draws attention to the communicator, which in this study represents *the researcher who produces an ETD*. Mersham's (1995:55) model proposes two main dimensions to the profile of the communicator as a starting point for analysis, namely socio-cultural circumstances and autobiographical circumstances.

Socio-cultural circumstances are a broad category that will impact on the researcher's ability to produce an ETD. Their exposure to ICTs in their educational background and access to facilities at home or at the research institution will all have an impact on the way the communicator prepares his message.

Autobiographical circumstances are those specific to the individual researcher. This study will not deal with this aspect in depth, except to acknowledge their importance by referring to autobiographical comments of

certain respondents to the survey. For example, an individual may be 'technophobic' and therefore find himself/herself at odds with the process of ETD production and management.

#### **2.2.4 ENCODING**

Encoding is a key to the success or failure of any message in the communication process. As Mersham (1999:190) points out, the choice of linguistic code ('language') is one of the first encoding decisions. Some institutions of higher education in South Africa currently have policies that allow for the publication of TDs in South African official languages other than English. The question may be asked "What is the impact of these policies on ETDs?" The present answer may well be that only people understanding that specific language would be able to understand the message, however in the future as mentioned before software could be developed to translate or decode the message into ones home language through for example directed navigation within the user interface.

Other levels of encoding ETDs include the use of a range of codes that ICTs make possible in encoding the message content of the research. These include visual codes, such as colour, animation, as well as audio codes such as speech, event and object sounds or a combination of both as in video.

Another category of ETD encoding is that of technical encoding. This includes the ability of the communicator to follow the coding requirements of using different software to produce different formats, e.g. .doc, PDF, html etc.

Another coding option is the use of metadata. Pockley (1991:223) describes metadata as resource description; the kind of information found in a library catalogue. He advises that quality metadata is easy to provide at the point of ETD creation but is usually difficult, expensive or impractical to add with hindsight. Pockley (1991:349) further advises that accurate metadata improves accuracy and increases the recall of the content of ETDs. This subject will be revisited in the next chapter when Hackos (2002:4) explains

how metadata forms the foundation of the information model for dynamic information delivery on the Web.

Navigation structures including hyperlinks, bookmarks, table of contents and indexes encoded into the body of work will lead to an efficient information model that provides the user with more than one way to get to parts of the study, other documents and web-based resources.

### **2.2.5 THE MESSAGE, SIGNS AND SYMBOLS**

The message is encoded and transmitted by the medium. When signs and symbols are arranged in conventionally agreed ways (codes) they form messages. The message in the context of disseminating academic research can be delivered in written or spoken language. TDs are normally submitted in written format with parts or summaries sometimes being verbally presented.

Scientific signs and symbols must also be considered; Dobratz (1991:168) gives the example of the text formatting system LaTeX which is used within the natural and engineering sciences to mark up mathematics and algorithmic graphics in scientific documents.

### **2.2.6 THE MEDIUM**

The medium is the message receptacle, the means by which a message is encoded or 'wrapped'. When you talk to somebody, the air is the medium and the speech is the message. When you buy a book, paper is the medium, the print is the message and the communicator is the author.

Applied to ETDs the recipient/s opens a connection to the electronic message via the Internet which is the medium. According to Mersham (2001:220) the Internet represents an international 'network of networks'. It allows millions of computers and other electronic devices like cell phones and laptops to communicate with each other around the world via telephone lines, digital leased lines, fiber-optic lines and satellites connections in an economical

easy-to-use way. Connected devices on the Internet may use a number of different physical media but all are governed by a network protocol on a means of communication and what resources are to be shared.

According to the Linux home networking website (2006) the Transmission Control Protocol / Internet Protocol (TCP/IP) is a universal standard suite of protocols used to provide connectivity between networked devices. It is part of the larger Open System Interconnection (OSI) model upon which most data communications is based. The Internet Protocol (IP) which is on the third layer of the OSI model (Network Layer) is responsible for routing messages using the best path available. The Transmission Control Protocol (TCP) which occurs on the fourth layer of the OSI model called the transport layer ensures properly sequenced and error free transmission. For manageability, the data is usually split into multiple packets each with its own error detection bytes in the header of the packet (Linux home networking website: 2006). The remote computer then receives the packets and reassembles the data and checks for errors, if there are none it then passes the data to the program that expects to receive it (Linux home networking website: 2006).

According to Bowers (2003) there is not a single aspect of communication that has not been significantly affected by the Internet, including the ease of accessing information, the lower cost of accessing information, the increase in interactions over space and time between people and the overall potential scale of communication.

### **2.2.7 THE RECIPIENT(S)**

This aspect of the model represents the audience of the message. Academics, researchers etc. accessing ETDs become recipients of their messages.

Mersham's (1995:55) model proposes the same two dimensions to the profile of the recipient as for the communicator, namely socio-cultural circumstances and autobiographical circumstances.

Socio-cultural circumstances include, for example, the culture in which the recipient experiences his or her daily life. For example, in Europe, North America and Australasia, broadband is culturally familiar and commonly experienced, which is not the case currently in South Africa. Socio-cultural circumstances also include the organizational culture of the institution to which the researcher belongs. Does it support ETDs or not? Socio-cultural circumstances impact on the recipients' ability to access and understand the message.

In the context of ETD access the requirements would be three fold namely:

1. Facilities to access the Internet.
2. The technical skills in the ability to use an Internet browser, search for and navigate to the message then download and open in the application it was intended for.
3. The language skills to decode the message.

Autobiographical circumstances are those specific to the recipients field of interests, life experiences, etc which allows recipients to understand messages associated with specific fields of research.

Recipients assume the interchangeable role of communicator in the model when they comment on or review the ETD message, thereby contributing to the valuable function of peer-reviewing within scholarly communication.

## **2.3 ASPECTS OF INFORMATION MODELING**

### **2.3.1 INTRODUCTION**

According to the Gilbane report (2002:1) ten years ago few institutions did anything with unstructured or 'document' information other than print it, or scan and save it. The report (2002:1) describes the early use of proprietary markup languages to build information models to better manage, reuse, and benefit from their non-relational information assets with tools such as Standard Generalized Markup Language (SGML). Today many organizations are using Extensible Markup Language (XML) technology to add structure and extended functionality to structured documents and data on the Web

(Gilbane report; 2002:1); this will be explained in more detail later in this chapter.

The report (2002:2) goes on to publish an excerpt from a book called: Content Management for Dynamic Web Delivery by JoAnn Hackos from which we will draw her explanation of why designing a carefully planned information model is a significant component of a content management strategy. We can assume that the same strategy would apply to ETDs.

### **2.3.2 WHAT IS AN INFORMATION MODEL**

According to Hackos (2002:2) an information model provides the structure for organizing your content so that it can be delivered and reused in a variety of innovative ways, which will enhance search and retrieval, making it possible for authors and users to find the information resources they need promptly and straightforwardly.

Hackos (2002:2) states that the information model is the ultimate content-management tool which requires analysis, careful planning, and a lot of feedback from the user community in its creation to ensure their points of view are taken into account. Hackos (2002:3) explains that the information model is an *organizational framework that is used to categorize information resources on which to base the publishing design for both print and electronic information delivery.*

### **2.3.3 THE THREE-TIERED STRUCTURE OF AN INFORMATION MODEL**

Hackos (2002:3) goes on to describe a three-tiered structure for the Information model. At base, the first tier consists of the metadata dimensions that identify how information will be categorized and labeled for both internal and external use. The second tier sorts information assets into information types. The third tier provides structure for each information type, outlining the content units that authors use to build the information types. Figure 2.3 below illustrates the three-tiered structure.

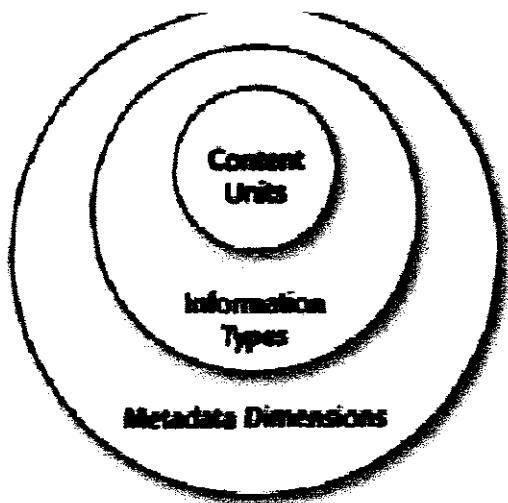


Figure 2.3 Hackos's Information Model from the Gilbane report (2002:4).

Hackos (2002:4) explains that the metadata dimensions which are identified as the foundation of the information model become the attributes and values of the metadata that will be used to label the messages in the archive.

According to a report by the Online Computer Library Center, Inc Working Group (2002:4), *preservation metadata is the information necessary to practically render digital resources over the long-term. Creating accurate preservation metadata will allow straightforward access to digital resources providing the understanding of the appropriate software applications designed to render the different archived resources.*

Hackos (2002:4) continues to explain that the information types should provide communicators with the starting point for creating well-structured messages that represent a particular purpose in communicating information and the content units will describe the content that is used to construct each information type. For example TDs have recommended structures for disseminating research in great detail while articles for journals have recommended structures for a more condensed dissemination of the same research; a table of contents in the TD describes the content units of the study.

#### **2.3.4 STATIC AND DYNAMIC INFORMATION MODELS**

Figure 2.4 shows the hierarchical view of folders in Windows Explorer. Hackos (2002:5) states that this static organization is useful as long as the

user understands the information and the organizational logic of the Information Model embodied in the hierarchical design.

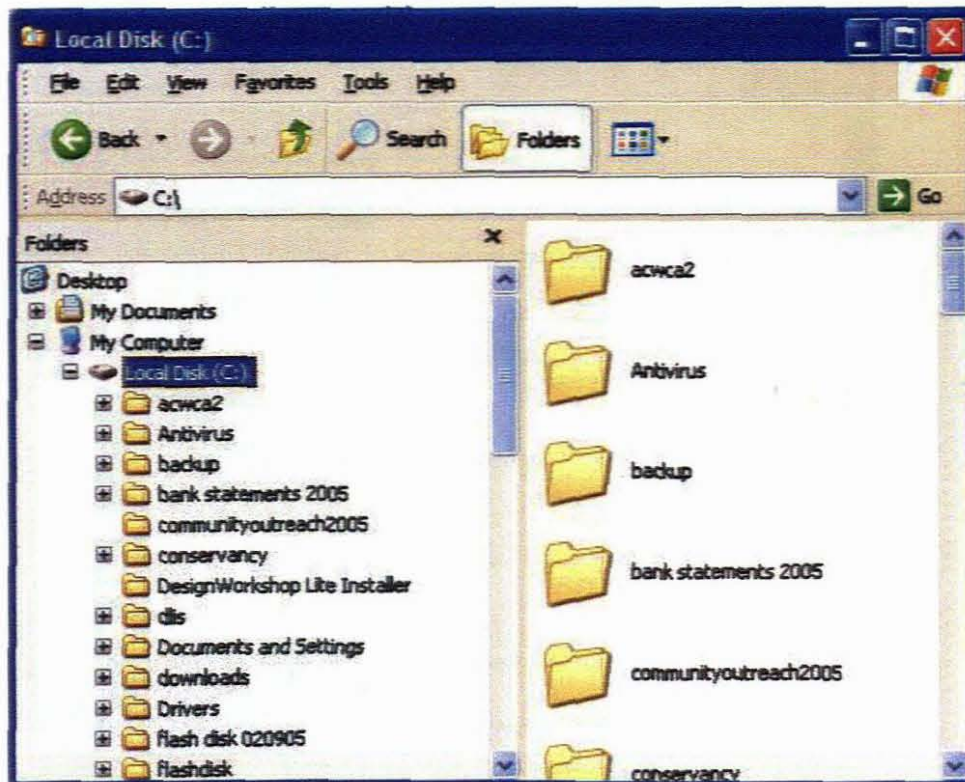


Figure 2.4 Hierarchical organisational structures in Windows Explorer.

Figure 2.5 (see Page 31) shows the dynamic search function within Windows Explorer, which allows the user to search for data using various criteria like data type, file names, when the data was last modified or a word or phrase within the document etc.

Hackos (2002:5) advises that in thinking up an Information Model for dynamic Web delivery, you need to avoid creating a static system where there is only one way to find a particular piece of information.

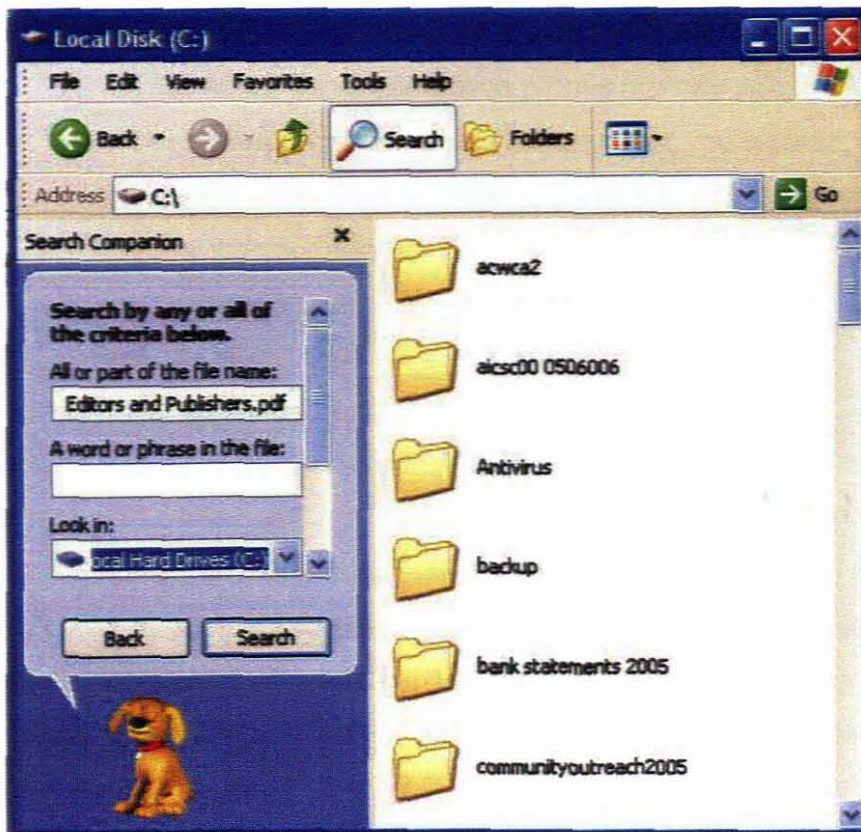


Figure 2.5 Dynamic search functions within Windows Explorer.

Hackos (2002:2) uses the following characteristics to identify Web sites with no underlying Information Model.

- You can't tell how to get from the home page to the information you're looking for.
- You click on a promising link and are unpleasantly surprised at what turns up.
- You keep drilling down into the information layer after layer until you realize you're getting farther away from your goal rather than closer.
- Every time you try to start over from the home page, you end up in the same wrong place.
- You scroll through a long alphabetic list of all the articles ever written on a particular subject with only the title to guide you.

Hackos (2002:2) then compares these characteristics to a Web site with a well-conceived Information model where:

- On the home page, you notice promising links right away.
- Two or three clicks get you to exactly what you wanted.
- The information seems designed just for you because someone has anticipated your needs.
- You can read a little or ask for more – the cross-references are in the right places.
- Right away you feel that you're on familiar ground – similar types of information start looking the same.

Coorough (2001:38) states that dynamic web sites provide information and offer some form of interactivity such as e-mail, searches, questionnaires, and order processing or credit card purchasing. In addition, Coorough (2001:38) states that users are often able to customize the site to meet their personal needs.

### **2.3.5 ORGANIZATIONAL STRUCTURE**

Coorough (2001:39) emphasises that there is a basic reason for organizing information and that is that our human brains like order. Part of planning a CMS web site and developing the site structure involves developing an effective organizational structure. Coorough (2001:40) goes on to explain that the organizational structure deals with how information is stored on the server and made available to the user while the navigational structure deals with how the user actually gets to the information. According to Coorough (2001:40) there are a variety of different ways in which to organize information.

#### **2.3.5.1 Hierarchical**

The hierarchical method organizes data using top-down approach for example folders a-z in Windows explorer. Coorough (2001:40) states that this is a common organizational structure because data can be easily and naturally broken down into sets and subsets that form the hierarchy. Coorough (2001:40) goes on to say that this method of organization is very common on

the Web because most Web sites start with a home page that provides general information that becomes increasingly specific.

Coorough (2001:40) differentiates between Hierarchical systems that are broad and shallow as illustrated in figure 2.6 (below), or narrow and deep as illustrated in figure 2.7 (below).

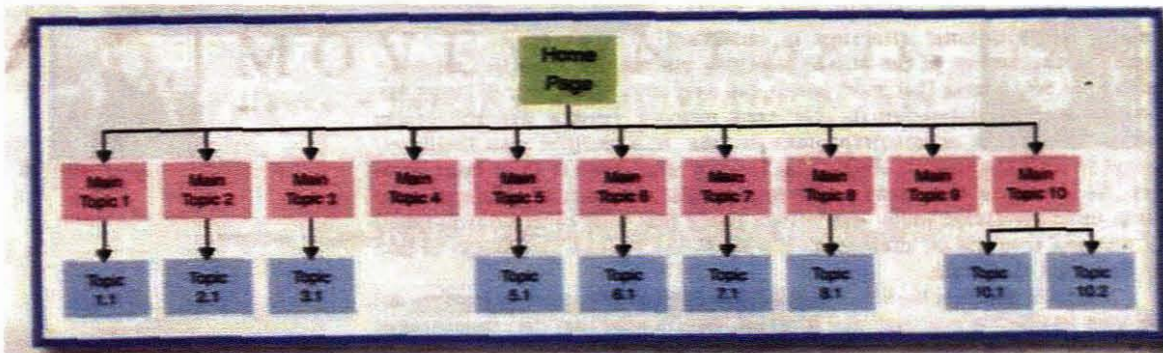


Figure 2.6 Broad and shallow Hierarchical system from Coorough (2001:40).

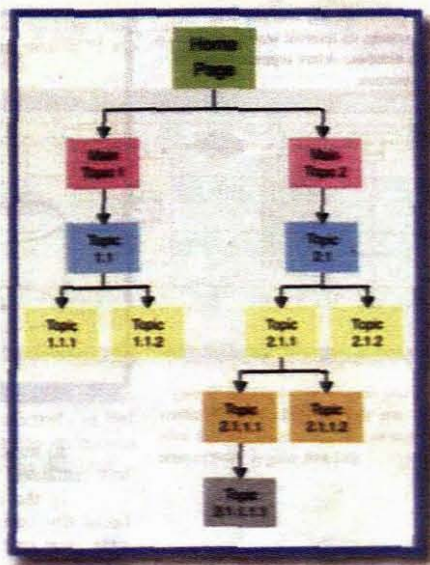


Figure 2.7 Narrow and deep Hierarchical system from Coorough (2001:41).

### 2.3.5.2 Nonlinear

Coorough (2001:40) explains that a nonlinear organizational structure has no prescribed or sequential path. Coorough (2001:40) states that most information on the World Wide Web consists of information that is read non-sequentially. In a nonlinear organizational structure the information is

organizes so that there are links from one separate piece of information to another as shown in figure 2.8 (below).

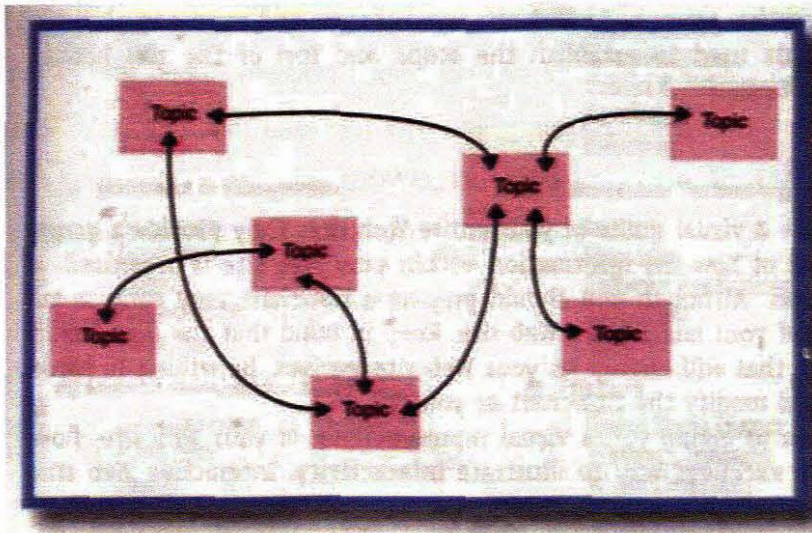


Figure 2.8 Nonlinear organizational structures from Coorough (2001:41).

#### 2.3.5.3 Linear

Coorough (2001:41) states that the linear method organizes information sequentially as illustrated in figure 2.9 (below). Coorough (2001:41) says that this is the traditional method of navigation used in Web pages that have a set of sequence that must be followed in order for the information to make sense. Coorough (2001:41) states that although this is the most common organizational structure used in most analogue media such as video and printed books or TDs, it is a very uncommon organizational system on the Web.

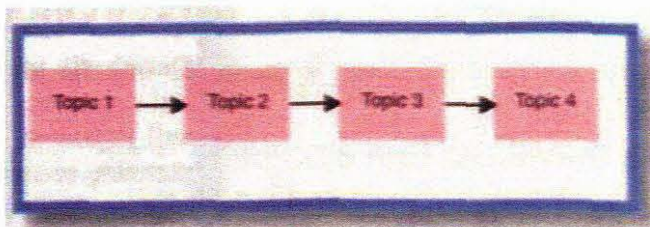


Figure 2.9 Linear organizational structures from Coorough (2001:41).

#### 2.3.5.4 Database-Driven

Coorough (2001:41) explains that the database-driven organizational structure is becoming increasingly popular because of the huge volumes of short bits of information found on the Web. Coorough (2001:41) states that any Web site that uses an internal search engine falls into this category and this organizational structure is very common on library and research type sites, as well as on e-commerce Web sites.

As illustrated in figure 2.10 (below), this organizational structure organizes data into containers that can be accessed upon request from the user.

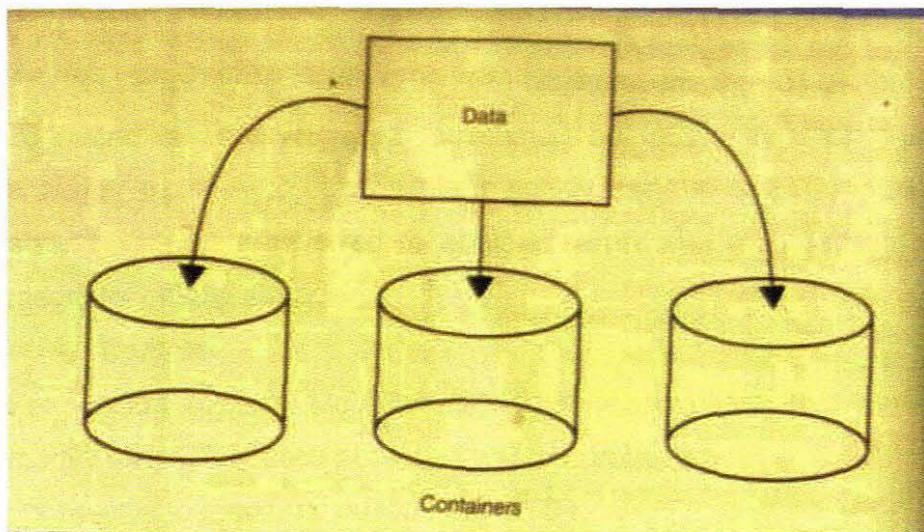


Figure 2.10 Database-driven organizational structures from Coorough (2001:41).

Coorough (2001:42) concludes that in the reality, most Web sites use a combination of these organizational structures.

#### 2.3.6 LIBRARIES AND ARCHIVING ETDs

McMillan (2001:66) calls the role of libraries to archive ETDs tenuous because not enough time has passed to prove that digital documents can live for decades in publicly accessible digital libraries. McMillan (2001:66) argues that libraries should establish ETD project Web sites that document the evolving

initiative, encourage active participation among interested parties, as well as providing submission forms, policy statements and ETD samples.

McMillan (2001:67) states that the best chance electronic information has of being preserved is when it is used online regularly and continually and three factors need to be addressed when archiving of ETDs namely:

### **1. Access**

McMillan (2001:68) states that the first goal is to have all ETDs online and available around-the-clock from a stable server and should this primary server's resources get depleted then statistics from the data-base should be used for migrating unused ETDs to a secondary server for archiving.

### **2. Security**

#### **• Data**

For data security McMillan (2001:68) recommends that ETDs that have been submitted but not yet approved should be frequently backed-up because changes are likely to occur while a weekly back up of ETDs in all directories should then be transferred to another server and recorded on a preferred media for off-line storage.

#### **• Content**

For content security McMillan (2001:68) recommends that authors cannot modify their ETDs once approved without proper approval. Readers must also not be able to modify or replace any ETDs.

### **3. Format Migration**

McMillan (2001: 69) states that the library should share with the university the responsibility to guarantee that ETDs will be available both within and outside the scholarly community *ad infinitum*. McMillan (1991:69) says to do this will mean migrating current formats to new standard formats not yet in existence while recommended standard formats should be the only acceptable files approved.

Table 2.1 in appendix 1 shows the recommended formats from Virginia Tech website (2006). These formats, it is suggested, should form the basis of planning for ETD projects.

### 2.3.7 METADATA

Pockley (2001:223) states that there is nothing new about the concept of metadata, which is resource description or the kind of information found in a library catalogue. What Pockley (2001:223) says is new in the digital world is the essential role that the creator of electronic documents now plays in providing this information. Pockley (2001:349) recommends that in addition to creating an ETD, students should become their own cataloguers creating accurate metadata, which improves accuracy and increases the recall of the content of ETDs. Pockley (2001: 349) warns that this requires some of the *basic skills of resource description and diligence in avoiding three language problems that cause poor precision namely:*

- **Polysemy:** words with multiple meanings. Pockley (2001: 349) gives the example of the word spring, which could mean the season, a physical coil or a source of water.
- **Synonymy:** different words with the same meaning. Pockley (2001: 349) gives the words 'ball,' 'sphere,' and 'orb,' or 'scuba diving' versus 'skin diving' as examples and states that good metadata should draw these materials together, despite their use of different synonyms.
- **Ambiguity:** Pockley (2001:350) says to overcome ambiguity a search engine must understand the meaning of a word and not just match the spelling, if it is going to make a distinction between different meanings of the same word.

Pavani (2001:296) states that the richer and more versatile the metadata model is the more resources it takes and for some situations it may be necessary to adopt the simplest possible model in order to make the metadata capture feasible. Pavani (2001:296-297) states that it is agreed that the Dublin Core Metadata Element Set (DCMES) should be the minimum classification scheme to be used for ETDs.

Penman and Milne (2004) state the Open Archives Initiative Protocol for Metadata Harvesting (OAI: PMH) provides an application-independent interoperability framework based on metadata harvesting. Penman and Milne

(2004) explain that OAI-compliance means all archives created in this way are "interoperable" as they use the same (OAI) convention for tagging their metadata (author, title, date, journal, etc.). Penman and Milne (2004) state that this means the contents of all such archives can be navigated and searched effortlessly, as if they were all in one worldwide "virtual" archive.

### **2.3.8 CATALOGUING**

Pavani and McMillan (2001:305) state that many traditional library systems exchange and store records using the Machine Readable Catalog Format (MARC), which has in the region of a thousand fields, many with subfields which can be repeated. This according to Pavani and McMillan (2001:305) allows a very detailed description of the items including a specific field (856) to identify electronic objects associated with the intellectual item and its other physical instances.

Pavani and McMillan (2001:305) portray the Dublin Core Metadata Element Set (DCMES) as a set of fifteen (15) attributes divided into three (3) groups: content, intellectual property and instantiation and associated to them there are the Dublin Core Qualifiers that enhance the identification of the items.

Pavani and McMillan (2001:305) then describe the Resource Description Framework (RDF) as the foundation for processing metadata as it specifies a representation for metadata as well as the syntax for encoding and transporting this metadata. According to Pavani and McMillan (2001:305) the objective of RDF is to yield interoperability of Web servers and clients, and to facilitate automation of processing of resources and it can be used to describe Web pages, sites or digital libraries.

Pavani and McMillan (2001:305) state that there is a relation between the MARC Format and the DCMES since there is an intersection between the two sets of attributes while DCMES can also be used with the RDF representation.

### **2.3.9 INFORMATION RETRIEVAL**

According to Pavani (2001:319) the next step after identification of the items of the digital library, or ETDs, is to address storage of the cataloguing attributes and the process of searching and retrieving. Pavani (2001:319) stresses that the quality of retrieval is reliant both on the programming of the search and retrieve functions but, as important as this, on the quality of the information used to catalog the items of the collected works.

Pavani (2001:319) recommends databases as suitable tools to store, search for and retrieve information. Pavani (2001:319) states that before implementing the database, the database model must be created after the metadata model has been defined and related to other existing identification procedures such as traditional cataloguing on an automated library system.

Pavani (2001:319) advises that if the traditional Online Public Access Catalogue (OPAC) is to be maintained during the ETD program the attributes that are present in the OPAC should not be duplicated and a link between the OPAC record and the ETD metadata be created.

Pavani (2001:319) states that no matter where the information is stored, the user or recipient should be able to perform the types of search that are standard in traditional library systems such as for author, title, keywords, subjects, ISBN, etc. Pavani (2001:319) also mentions that this information that is used to identify the items is language dependent (title, keywords, subjects, etc) and if the database holds only one language per record, search procedures are to be performed using the arguments in this language. However if a multilingual database is modeled, Pavani (2001:319) recommends that the search be language independent, i.e., that the argument be checked against all languages while all its language instances should be displayed to the user so that he/she can retrieve the information in the language of their choice.

### **2.3.10 FULL TEXT SEARCHING**

Fox (2001:351) states that when all of the text of an ETD is available for searching, a digital library system is said to support full text searching where

users can submit queries that call for documents that have particular phrases, words, categories, or word stem. Fox (2001:351) says that for full text searching to work, the entire document must be analyzed, and used to build an index that will speed up searching but require a good deal of storage space, Fox (2001:351) estimates around thirty percent (30%) of the size of the texts themselves.

Fox (2001:351) warns that full text searching may lead to decreased precision, which can be compared to Internet searches today where many links to documents that only makes casual mention of a keyword, when the majority of the information is about other irrelevant topics.

Fox (2001:351) however points out that full text searching may improve recall, since works can be found that are not classified to be about a certain topic and full text searching often yields relevant paragraphs, rather than just a pointer to a document that then must be scanned to establish relevance.

### **2.3.11 THE EXTENSIBLE MARKUP LANGUAGE (XML)**

Dobratz (2001:352) states that the Extensible Markup Language (XML) is the universal format for structured documents and data on the Web. Dobratz (2001:353) explains that before XML there was SGML, developed in the early '80s and Hypertext Markup Language (HTML), whose development started in 1990. Dobratz (2001:353) points out that the designers of XML simply took the best aspects of SGML, guided by their experience with HTML, and produced something that is no less powerful than SGML, but vastly more standard and simpler to use.

Dobratz (2001:353) states that the main philosophy of SGML and XML is the strict separation of content, structure and layout of documents and that most ETD projects use either the SGML standard or the definition of the World Wide Web Consortium (W3C) XML 1.0.

Patel and Fox (2001: 354) state that XML is a method for putting structured data into a text file format and say that "structured data" like spreadsheets, address books, configuration parameters, financial transactions and technical drawings result from software applications which often also store it on disk,

these applications can either use a binary format or a text format where the latter allows one, if necessary, to look at the data without the program that produced it. Patel and Fox (2001: 354) portray XML as a set of rules for designing text formats for such data, in a way that produces files that are easy to produce and process by computers, that are clear-cut, and that avoid common drawbacks, such as lack of extensibility, lack of support for internationalization / localization, and platform dependency.

Patel and Fox (2001: 355) believe that XML documents can be preserved for a long time because XML is a simple data format. Patel and Fox (2001: 355) also state that it can be written in one hundred percent (100%) pure ASCII text as well as in a few other well-defined formats with the ASCII text reasonably resistant to corruption.

Patel and Fox (2001: 355) say that XML documents have four-stage life cycle:

- Created using a basic text editor like notepad or a 'what you see is what you get' WYSIWYG editor like Macromedia Dreamweaver.
- The XML parser reads the document and converts it into a tree of elements.
- The parser passes the tree to the browser
- Browser displays it.

Patel and Fox (2001: 355) say that it is important to know that all these processes are independent from each other.

Patel and Fox (2001: 355) state however that XML comes with its own bag of discomforts namely:

- Conversion from word processing forms to XML requires more planning in advance, different tools and broader learning about processing concepts than for PDF therefore converting may be complicated, difficult and time consuming.
- Writing directly in XML by using XML authoring tools requires some prior knowledge of XML.

- XML is very strict regarding the naming and ordering of tags. It is also case sensitive illustrating the relative effort required by students to prepare ETDs in this form.

### **2.3.12 MULTIMEDIA**

Coorough (2001:49) states that multimedia refers to the creation of a multi-sensory experience to convey messages and ideas. Coorough (2001:49) says that to create this multi-sensory experience, different multimedia elements are used including text, graphics, animation, sounds, and video. Coorough (2001:49) acknowledges that it takes just the right combination of elements to effectively portray and illustrate a message.

Jacques (1997) in Thornhill *et al* (2002:11) observes that images are a source of knowledge, persuasion and pleasure every bit as powerful as the written word. Thornhill *et al* (2002:11) note several researchers (e.g. Shepard and Cooper, 1982, Mayer and Gallini, 1990) who made the connection between visual clues, the memory process and the recall of new knowledge. Thornhill *et al* (2002:11) highlight that the way in which sound and image can complement each other in portraying the message has also been well researched. Goodyear and Steeples (1998) in Thornhill *et al* (2002:11) state that video can provide vivid descriptions to communicate tacit information and knowledge hard to describe through text.

Coorough (2001:50) states that a well designed multimedia Web pages / sites appeals to more than one sense as it is important to give users enough stimulation to grab their interest without overwhelming them. Coorough (2001:50) concludes that balance is the key and if achieved, retention and understanding are enhanced while if the balance is off, users will either become bored or confused.

Thornhill *et al* (2002:13) remind us of Paivio's (1971, 1991) studies and the dual-code theory where he concluded that information that is supported by both aural and visual cues should increase recall and retention of that

information. Video, text and audio put together can debatably enhance a communication channel, and the dissemination of TDs.

Fox (2001:370) states that while most ETD collections today allow search for multimedia content based on descriptions of such content (i.e., metadata), in the future, as collections develop and have richer collection of multimedia elements, it is likely that multimedia databases and multimedia content search software like Google's picture search will be more widely used to search and retrieve multimedia content.

### **2.3.13 SOFTWARE OPTIONS FOR HOSTING ETDs AND PROVIDING THE USER INTERFACE**

The use of web interfaces is the logical choice to provide the user interface to ETD archives because of the benefits of using the Internet as the medium for the message and the ability to apply a dynamic information model to the interface.

Penman and Milne (2004) list four open source options namely:

ETD-db developed at Virginia Tech, EPrints developed at the University of Southampton, DSpace developed jointly by MIT Libraries and Hewlett-Packard and Greenstone Digital Library produced by the New Zealand Digital Library Project at the University of Waikato.

Penman and Milne's (2004) advise that the selection criteria for choosing the appropriate software should consider the following key factors:

- **Open Source Software** means that the acquisition cost is negligible as the source code is freely downloadable with no license fee/s needed. Also the creation of a community driven approach helps identify features that are useful and most appropriate to the users needs. Furthermore the ability to alter code and further develop functionality tailored for individual institutions and in turn share these updates and additions.
- **Functionality of the Software** where the Open Archives Initiative Protocol for Metadata Harvesting (OAI: PMH) provides an application-

independent interoperability framework based on metadata harvesting, which allows contents of all such archives to be searched effortlessly worldwide. The software must also provide the ability for document management and storage capability as well as searching and browsing abilities.

- **Sustainability Aspects of the Software** entail an established user base to share institutional experiences and knowledge as well as help overcome initial barriers including hardware / software requirements / versions and installation / implementation issues. Support offered through mailing list to allow this knowledge to be shared and used and online guidance from the creator in the form of concise documentation, frequently asked questions and bug fixes etc.

Lastly they should give continued support and work on the core from the original creator to grow and develop the software.

Penman and Milne (2004) then state that DSpace and EPrints both have a proven record at successfully creating digital libraries for the capturing, storing, indexing, preserving, and distributing a huge variety of document types and content including conference papers, articles and reprints, working papers, images: diagrams / photographs, audio files / video files and scanned and reformatted older (paper based) documents.

## **2.4 SUMMARY**

In this chapter I have highlighted the perceived benefits and controversies surrounding ETDs. These perceptions will be tested by specific questions in the survey to see how valid they are in a South African academic setting. I have then applied a communication and information model to the ETD lifecycle to heuristically explain how attributes of ETDs influence the creation and dissemination of this electronic media. I have ended the chapter by discussing possible database driven software applications that can archive ETDs using metadata to catalogue a digital repository, which becomes accessible through a web interface.

## **CHAPTER THREE:        METHODOLOGY**

### **3.1            INTRODUCTION**

Leedy (1989:140) states that we learn some truth by observing the events taking place in the world around us and that the descriptive survey or normative survey method of research describes exactly what the researcher concludes after examining survey data. According to Leedy (1989:140) the name normative implies the assumption that what ever we observe in the present moment is normal and that under the same conditions in the future could be observed again. Leedy (1989:140) explains that the basic assumption underlying such an approach is that given phenomena follow a common pattern, or norm. Leedy (1989:141) however reminds us of the wisdom of Heraclitus who stated: 'There is nothing permanent but change'. Norms today will not be the norms in the future especially when it comes to the ever changing world of ICTs and the different ways we can make use of them.

Leedy (1989:141) explains that the meaning of survey is to "look or see over or beyond" but emphasizes that in research we have many ways of seeing that have nothing to do with physical vision. According to Shannon *et al* (2002) the Internet has had great impact on the field of survey research and the number of electronically-administered surveys continues to grow.

Myers (1997:2) introduces some philosophical perspectives when he states: "All research (whether quantitative or qualitative) is based on some underlying assumptions about what constitutes 'valid' research and which research methods are appropriate". He goes on to state that most pertinent philosophical assumptions are those which relate to the underlying epistemology which guides the research, bearing in mind that epistemology refers to the assumptions about knowledge and how it can be obtained.

Orlikowski and Baroudi (1991), following Chua (1986) in Myers (1997:3) suggest three categories for qualitative research, based on the underlying research epistemology: positivist, interpretive and critical.

Myers (1997:4) explains that interpretive researchers start out with the assumption that access to reality either given or socially constructed is only through social constructions such as language, consciousness and shared meanings. This study's approach puts emphasis on shared meanings of academics, postgraduate students, librarians and information technology specialists working and studying in South African. The research conducted also fits the broad assumptions of the social constructivist approach where the 'meaning' and therefore 'values' of ETDs are not an empirical given but something that is negotiated through debate and interpretation by the respondents themselves as they 'construct' their concept reality of ETDs while participating in the study.

### **3.2 OBJECTIVES**

The following questions arise and form the basis of this study:

- Which South African tertiary institutions are participating in this worldwide trend of disseminating ETDs and what is the nature and extent of this participation?
- What are the benefits and controversies associated with ETDs?
- What are the perceptions of academics toward this method of TD dissemination?
- How can communication and information models be used heuristically to explain the ETD lifecycle?

### **3.3 THE SURVEY INSTRUMENT**

An electronic survey (see <http://www.lis.uzulu.ac.za/etd/survey/index.html>) was posted on the World Wide Web (WWW). The questions were chosen to solicit the respondents demographic information, their academic field of research and experience, their sources of reference or research material, the software being used in their recording, analysis and presentation of their postgraduate work and or research, their general perceptions and

reservations of ETDs and whether their institutions accepted this format of TD. The last part of the survey allowed the respondents to optionally give personal information, add additional comments and to request to be notified once the results of the study were published.

The survey instrument employed thirty-one questions and a variety of response options including drop down menus, check boxes, and radio buttons for the close-ended responses that would give the study its quantitative data. Text boxes placed next to many of the questions for optional open-ended feedback provided valuable qualitative responses from the participants.

### **3.4 PROCEDURE**

The participants were identified by published email addresses on their tertiary institutions or departmental websites. Email lists were compiled in Microsoft Excel and saved in a comma delimited format then imported into Outlook Express's address book. A total of six thousand one hundred and sixty six (6166) email addresses were obtained and added to the survey's address book. From the 17th of October 2005 to the 7th of November 2005 potential respondents were sent an e-mail message with a cover letter requesting their participation and a link to the survey instrument which was published to the following URL: <http://www.lis.uzulu.ac.za/etd/survey/index.html>. From the total of six thousand one hundred and sixty six (6166) emails sent out one thousand two hundred and eighty five (1285) were returned undelivered and four hundred and twenty two (422) responses were received of those twenty one (21) were duplicate submissions. After subtracting these undelivered emails (1285) from the overall sample (6166), and the duplicates from the overall submissions, a response rate of eight percent (8.2%) was obtained (i.e. 401 out of 4881).

### **3.5 SAMPLE**

According to Leedy (1989:152) sampling is appropriate whenever large populations that have an outward resemblance of homogeneity are to be investigated. For this study the population under consideration worked or studied at a tertiary institution. Therefore the outward resemblance of

homogeneity of the study would be people working or studying at a tertiary institution in South Africa with or in the process of obtaining a postgraduate education.

Leedy (1989:152) states that sampling can be divided into two major categories: nonprobability sampling and probability sampling. Leedy (1989:152) warns that in nonprobability sampling there is no way of predicting, estimating or guaranteeing that each element in the population will be represented in the sample. This study recognises that due to nonprobability convenience sampling a correlation in some cases between the recipients' interest and awareness in ETDs and their willingness to participate in the survey could introduce bias in support of ETDs into the results.

In the case of this study the sample included 'intended', or targeted recipients who were emailed and requested to undertake the survey, and untargeted recipients who were shown the study by one of the targeted recipients also known as snowball sampling.

According to Hill (1998), the calculation of an appropriate sample size generally depends on the size of the population in question. He states however that researchers using e-surveys generally cannot find out, nor even estimate the size of the population they are studying (Hill; 1998). Krejcie & Morgan (1970) in Hill (1998) calculate that as the population increases, the sample size increases at a diminishing rate and eventually remains constant at a point above three hundred and eighty (380) cases, thus they state that there is little to merit the cost and energy of sampling above this figure. According to Hill (1998), Alreck and Settle (1995) provide similar evidence.

### **3.6 CONCLUSION**

Using a descriptive survey method most objectives of the study were achieved and presented in the former and latter chapters of the study. Limitations in sampling could minimise the value of this research and introduce bias into the study.

## **CHAPTER FOUR: RESULTS AND DATA INTERPRETATION**

### **4.1 INTRODUCTION**

This chapter deals with the presentation and interpretation of the quantitative and qualitative data collected via an electronic survey. The target audience consisted of academics, postgraduate students, librarians and information technology specialists working and studying in South African tertiary education institutions.

### **4.2 RESULTS AND DATA INTERPRETATION**

The study uses frequency and descriptive statistics to summarize survey data obtained.

#### **4.2.1 SAMPLE**

From table 4.1 it can be seen that four hundred and one (401) participants took part in this survey from twenty two (22) different tertiary institutions in South Africa. As seen from the occupation frequencies in table 4.2 most participants selected an occupation of faculty member (51%), then researcher (20%), then post-grad student (17%), then librarian (5%), then IT specialist (4%) and the remaining twenty or four percent (4%) originated from administrative posts to positions in the Department of Education. The multi-taskers ability to select more than one occupation leads to some participants answering the questionnaire in more than one capacity.

Figure 4.1 shows that seventy six percent (76.6%) of three hundred and ninety eight (398) participants that answered question three had post-graduate degrees while four percent (4.3%) did not. Nineteen percent (19.1%) or 76 participants were still in the process of obtaining their post-graduate degrees.

Table 4.1 Tertiary Institutions where participants of the electronic survey worked or studied.

|         |                                                         | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|---------------------------------------------------------|-----------|---------|---------------|--------------------|
| Valid   | Cape Peninsula University of Technology (CPUT)          | 11        | 2.7     | 2.8           | 2.8                |
|         | Central University of Technology (CUT)                  | 8         | 2.0     | 2.0           | 4.8                |
|         | Durban Institute of Technology (DIT)                    | 17        | 4.2     | 4.3           | 9.1                |
|         | Monash University                                       | 1         | .2      | .3            | 9.4                |
|         | Nelson Mandela Metropolitan University                  | 6         | 1.5     | 1.5           | 10.9               |
|         | North-West University (NWU)                             | 23        | 5.7     | 5.8           | 16.8               |
|         | Pretoria University (Tuks)                              | 14        | 3.5     | 3.6           | 20.3               |
|         | Rhodes University                                       | 25        | 6.2     | 6.3           | 26.6               |
|         | Tshwane University of Technology (TUT)                  | 10        | 2.5     | 2.5           | 29.2               |
|         | University of Cape Town (UCT)                           | 33        | 8.2     | 8.4           | 37.6               |
|         | University of Fort Hare (UFH)                           | 1         | .2      | .3            | 37.8               |
|         | University of Johannesburg                              | 31        | 7.7     | 7.9           | 45.7               |
|         | University of KwaZulu-Natal (UKZN)                      | 14        | 3.5     | 3.6           | 49.2               |
|         | University of Limpopo (UL)                              | 24        | 6.0     | 6.1           | 55.3               |
|         | University of the Free State (UFS)                      | 58        | 14.5    | 14.7          | 70.1               |
|         | University of South Africa (UNISA)                      | 26        | 6.5     | 6.6           | 76.6               |
|         | University of Stellenbosch (US)                         | 5         | 1.2     | 1.3           | 77.9               |
|         | University of Venda for Science and Technology (UNIVEN) | 4         | 1.0     | 1.0           | 78.9               |
|         | University of the Western Cape (UWC)                    | 15        | 3.7     | 3.8           | 82.7               |
|         | University of Witwatersrand (Wits)                      | 30        | 7.5     | 7.6           | 90.4               |
|         | University of Zululand                                  | 32        | 8.0     | 8.1           | 98.5               |
|         | Walter Sisulu University (WSU)                          | 6         | 1.5     | 1.5           | 100.0              |
|         | Total                                                   | 394       | 98.3    | 100.0         |                    |
| Missing | No answer                                               | 3         | .7      |               |                    |
|         | Other                                                   | 4         | 1.0     |               |                    |
|         | Total                                                   | 7         | 1.7     |               |                    |
| Total   |                                                         | 401       | 100.0   |               |                    |

Table 4.2 Occupation Frequencies.

|          |                          | Responses |         | Percent of Cases |
|----------|--------------------------|-----------|---------|------------------|
|          |                          | N         | Percent |                  |
| Capacity | Q2op1: Faculty           | 293       | 51.2%   | 73.6%            |
|          | Q2op2: Researcher        | 112       | 19.6%   | 28.1%            |
|          | Q2op3: Post grad student | 95        | 16.6%   | 23.9%            |
|          | Q2op4: Librarian         | 30        | 5.2%    | 7.5%             |
|          | Q2op5: IT specialist     | 22        | 3.8%    | 5.5%             |
|          | Q2op6: Other             | 20        | 3.5%    | 5.0%             |
| Total    |                          | 572       | 100.0%  | 143.7%           |

Q3: Do you have a post-graduate degree?

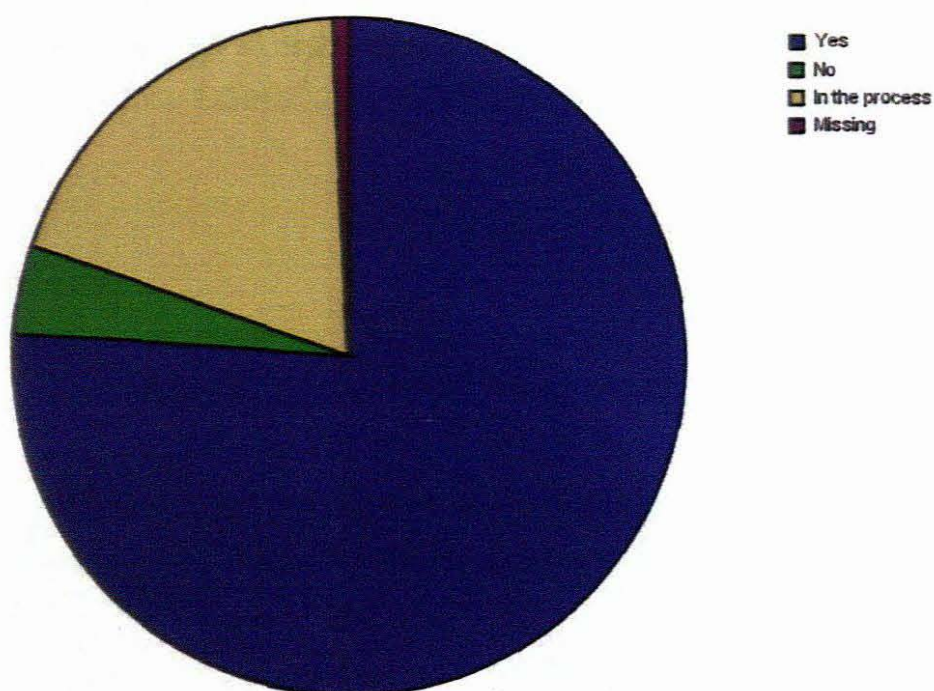


Figure 4.1 Tertiary education levels of participants.

#### 4.2.2 DEMOGRAPHICS

Ninety percent (90%) of the participants were South African, while about ten percent (9.5%) were from other countries.

Table 4.3 shows that fifty one percent (51%) of the respondents were male and forty eight percent (48%) female.

Table 4.3 Gender of participants.

|         |           | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|-----------|---------|---------------|--------------------|
| Valid   | Male      | 205       | 51.1    | 51.4          | 51.4               |
|         | Female    | 194       | 48.4    | 48.6          | 100.0              |
|         | Total     | 399       | 99.5    | 100.0         |                    |
| Missing | No answer | 2         | .5      |               |                    |
| Total   |           | 401       | 100.0   |               |                    |

Figure 4.2 shows that sixty eight percent (68%) of the participants were White, eighteen percent (18%) Black, six percent (6%) Coloured, four percent (4%) Indian and two percent (2%) Asian. Forty four percent (44%) stated that their home language was English while thirty seven (37%) percent said it was Afrikaans, sixteen percent (16%) was made up of a combination of IsiNdebele (0.7%), IsiXhosa (2.7%), IsiZulu (4.2%), Sepedi (2.7%), Sesotho (2.0%), Setswana (1.2%), SiSwati (0.5%), Tshivenda (1.5%) and Xitsonga (0.5%).

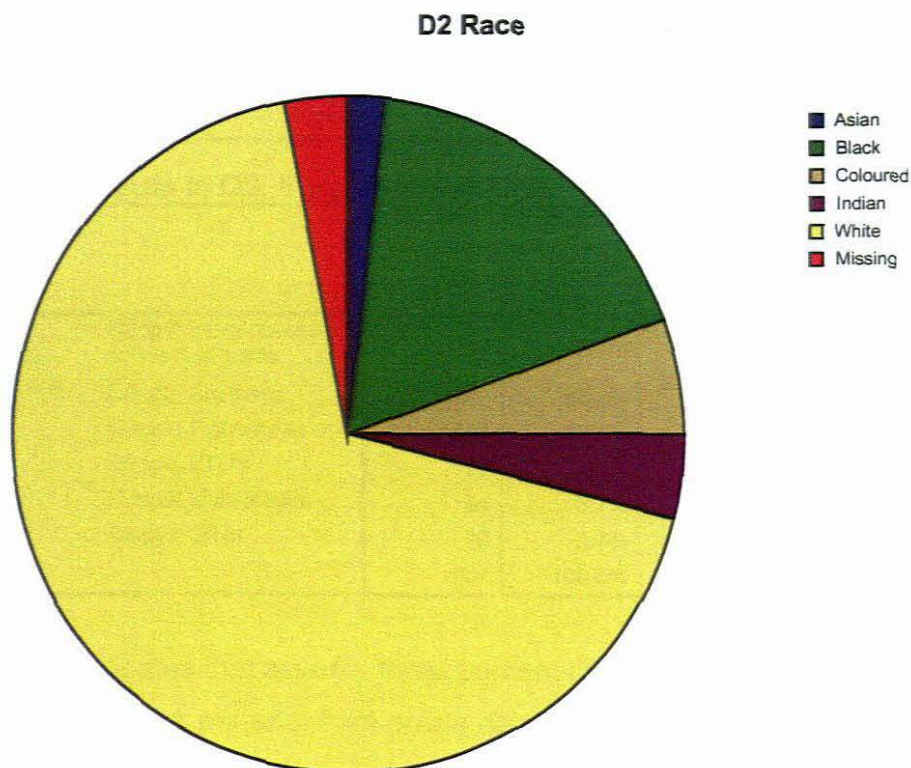


Figure 4.2 The race of the participants

Thirty one (31%) percent of the participants stated that they where in their thirties (30-39 years old) while twenty three percent (23%) where in their forties (40-49 years old), twenty one percent (21%) where in their fifties (50-59 years old), seventeen percent (17%) where in their twenties (20-29 years old), two percent in their sixties (60-69 years old) and one participant over sixty nine.

### 4.2.3 RESEARCH SOURCES

Table 4.4 shows that most participants have made use of online databases to obtain research material. Table 4.5 shows the response to Q9 which was a multiple response question and indicates what kind of material was downloaded from these online databases.

Table 4.4 Results to Q8. Have you used online databases to obtain research material?

|         |           | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|-----------|---------|---------------|--------------------|
| Valid   | Yes       | 369       | 92.0    | 92.7          | 92.7               |
|         | No        | 29        | 7.2     | 7.3           | 100.0              |
|         | Total     | 398       | 99.3    | 100.0         |                    |
| Missing | No answer | 3         | .7      |               |                    |
| Total   |           | 401       | 100.0   |               |                    |

Table 4.5 Results to Q9. If you have, what did you access/download?

|                           |                                       | Responses |         | Percent of Cases |
|---------------------------|---------------------------------------|-----------|---------|------------------|
|                           |                                       | N         | Percent |                  |
| Online Database Downloads | Q9op1: Have not used online databases | 17        | 2.1%    | 4.5%             |
|                           | Q9op2: Abstracts                      | 282       | 34.9%   | 74.4%            |
|                           | Q9op3: Full articles                  | 351       | 43.5%   | 92.6%            |
|                           | Q9op4: ETD's                          | 73        | 9.0%    | 19.3%            |
|                           | Q9op5: Multi-media                    | 54        | 6.7%    | 14.2%            |
|                           | Q9op6: Other                          | 30        | 3.7%    | 7.9%             |
| Total                     |                                       | 807       | 100.0%  | 212.9%           |

Table 4.6 indicates that seventy three percent (73%) of the participants found obtaining research material from online databases more efficient than from paperbound journals. Seventy three percent (73%) found online databases

search functions easy to use while eighteen (18%) said they found them difficult.

Table 4.6 shows the response to Q11: Do you find online databases more efficient to obtain research literature than paperbound journals?

|         |            | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|------------|-----------|---------|---------------|--------------------|
| Valid   | Yes        | 296       | 73.8    | 78.3          | 78.3               |
|         | No         | 46        | 11.5    | 12.2          | 90.5               |
|         | Don't know | 36        | 9.0     | 9.5           | 100.0              |
|         | Total      | 378       | 94.3    | 100.0         |                    |
| Missing | NAP        | 3         | .7      |               |                    |
|         | No answer  | 20        | 5.0     |               |                    |
|         | Total      | 23        | 5.7     |               |                    |
| Total   |            | 401       | 100.0   |               |                    |

Almost the entire sample (98%) had used internet searches to find research material. With eighty nine percent (89%) finding this method of searching for material efficient. As seen in figure 4.3 sixty eight percent (68%) said they found relevant material most of the time while eighteen percent (18%) said they found the material all the time. Ten percent (10%) said they seldom found relevant material and only three percent (3%) of the participants said they never found relevant material.

**Q23 How often do you find relevant material on the Internet?**

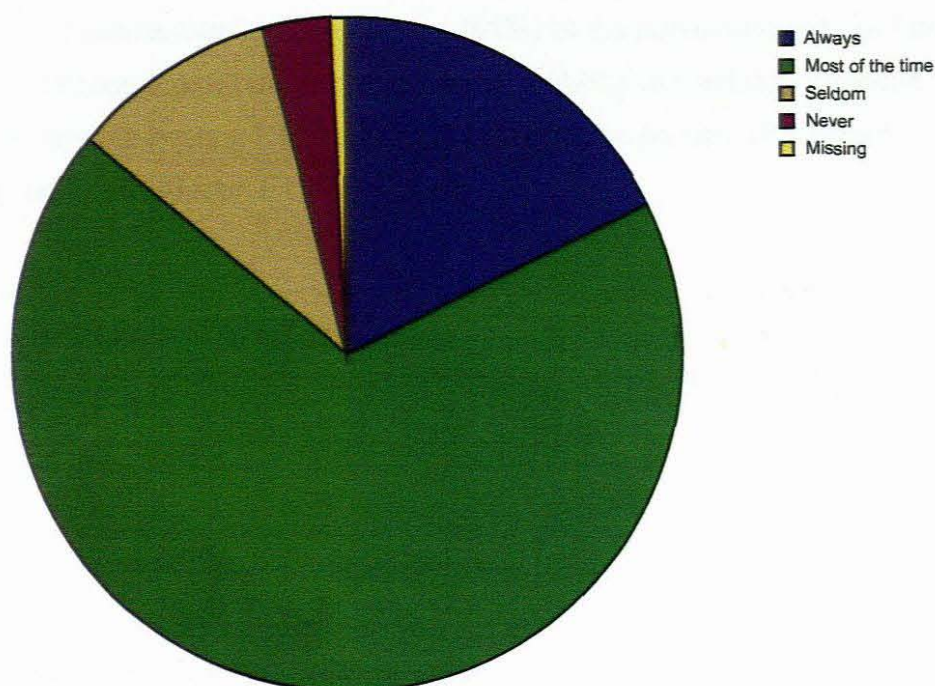


Figure 4.3 Results of Q23.

Table 4.7 tabulates the results of question twenty four which was a multiple response question. It shows that sixty three percent (63%) of the sample indicated that their sources were online which included online journals (25%), online databases (22%) and digital libraries (16%). Thirty seven percent (37%) indicated that they relied on paperbound material from libraries (19%) and journals (18%).

Table 4.7 Sources of research material.

|                              |                                          | Responses |         | Percent of Cases |
|------------------------------|------------------------------------------|-----------|---------|------------------|
|                              |                                          | N         | Percent |                  |
| Sources of research material | Q24op1: Libraries (Paperbound material)  | 184       | 19.2%   | 53.8%            |
|                              | Q24op2: Digital Libraries (Include ETDs) | 153       | 15.9%   | 44.7%            |
|                              | Q24op3: Journals (Paperbound material)   | 168       | 17.5%   | 49.1%            |
|                              | Q24op4: Journals (Online material)       | 244       | 25.4%   | 71.3%            |
|                              | Q24op5: Online databases                 | 211       | 22.0%   | 61.7%            |
| Total                        |                                          | 960       | 100.0%  | 280.7%           |

#### 4.2.4 PERCEPTIONS OF ETDs

Table 4.8 shows that fifty one percent (51%) of the participants were familiar with the concept of ETDs. Forty two percent (42%) had actually browsed digital libraries hosting ETDs and only twenty three percent (23%) had browsed South African ETD collections.

Table 4.8 shows the response to Q12: Are you familiar with electronic theses and dissertations (ETDs)?

|         |           | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|-----------|---------|---------------|--------------------|
| Valid   | Yes       | 206       | 51.4    | 51.8          | 51.8               |
|         | No        | 192       | 47.9    | 48.2          | 100.0              |
|         | Total     | 398       | 99.3    | 100.0         |                    |
| Missing | No answer | 3         | .7      |               |                    |
| Total   |           | 401       | 100.0   |               |                    |

Only twenty nine percent (29%) of participants said their tertiary institutions accepted ETDs however as seen in table 4.9 forty eight percent (48%) of participants thought that ETDs should be introduced at their tertiary institutions.

Table 4.9 shows the response to Q16: If your institution does not accept ETDs, do you think it should?

|         |            | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|------------|-----------|---------|---------------|--------------------|
| Valid   | Yes        | 191       | 47.6    | 60.8          | 60.8               |
|         | No         | 25        | 6.2     | 8.0           | 68.8               |
|         | No opinion | 98        | 24.4    | 31.2          | 100.0              |
|         | Total      | 314       | 78.3    | 100.0         |                    |
| Missing | NAP        | 69        | 17.2    |               |                    |
|         | No answer  | 18        | 4.5     |               |                    |
|         | Total      | 87        | 21.7    |               |                    |
| Total   |            | 401       | 100.0   |               |                    |

Looking at table 4.10 (see page 57) forty percent (40%) said that they thought that by adding multi-media content to TDs it would improve the expression of Graduate Education, while a further forty five (45%) answered that it would in

some cases. Eleven percent (11%) thought it would not improve the expression of Graduate Education.

Table 4.10 shows the response to Q25: Do you think by adding multi-media content (graphics, video, sound etc.) to theses and dissertations it will improve the expression of Graduate Education?

|         |               | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|---------------|-----------|---------|---------------|--------------------|
| Valid   | Yes           | 159       | 39.7    | 39.8          | 39.8               |
|         | No            | 43        | 10.7    | 10.8          | 50.6               |
|         | In some cases | 182       | 45.4    | 45.6          | 96.2               |
|         | No opinion    | 15        | 3.7     | 3.8           | 100.0              |
|         | Total         | 399       | 99.5    | 100.0         |                    |
| Missing | No answer     | 2         | .5      |               |                    |
| Total   |               | 401       | 100.0   |               |                    |

Eighty seven percent (87%) of the sample agreed that by adding theses and dissertations to an online database it would improve the accessibility and use of this research. While as seen from table 4.11 below sixty five (65%) indicated that they would not have any reservations on adding their thesis or dissertation to an online database. Thirteen percent (13%) said they would have reservations while twenty percent (20%) said they would in some cases have reservations.

Table 4.11 shows the response to Q27: Would you as an author have any reservations on adding your thesis or dissertation to an online database which everybody can access?

|         |               | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|---------------|-----------|---------|---------------|--------------------|
| Valid   | Yes           | 52        | 13.0    | 13.0          | 13.0               |
|         | No            | 261       | 65.1    | 65.4          | 78.4               |
|         | In some cases | 79        | 19.7    | 19.8          | 98.2               |
|         | No opinion    | 7         | 1.7     | 1.8           | 100.0              |
|         | Total         | 399       | 99.5    | 100.0         |                    |
| Missing | No answer     | 2         | .5      |               |                    |
| Total   |               | 401       | 100.0   |               |                    |

From table 4.12 it can be seen what the participants concerns where in publishing their thesis or dissertation (TD) electronically for everybody to access.

Table 4.12 Concerns about publishing a thesis or dissertation (TD) on the Internet for everybody to access.

|                                             |                           | Responses |         | Percent of Cases |
|---------------------------------------------|---------------------------|-----------|---------|------------------|
|                                             |                           | N         | Percent |                  |
| Concerns with publishing to the Internet(a) | Q28op1: No concerns       | 127       | 22.0%   | 32.3%            |
|                                             | Q28op3: Plagiarism        | 215       | 37.3%   | 54.7%            |
|                                             | Q28op4: Prior publication | 177       | 30.7%   | 45.0%            |
|                                             | Q28op2: Patent pending    | 58        | 10.1%   | 14.8%            |
| Total                                       |                           | 577       | 100.0%  | 146.8%           |

Figure 4.4 shows that sixty three percent (63%) of the participants thought that there would be cost benefits in publishing and storing electronic theses and dissertations (ETDs) compared to paper bound theses and dissertations (TDs).

**Q29: Do you think that there would be cost benefits in publishing and storing electronic theses and dissertations (ETDs) compared to paper bound theses and dissertations (TDs)?**

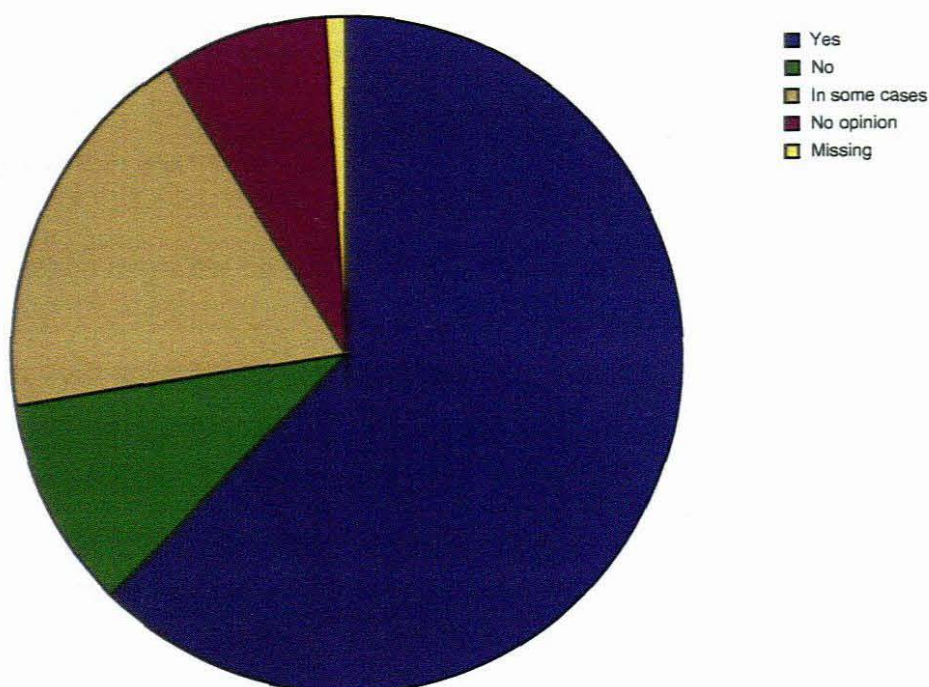


Figure 4.4 Results of Q29.

#### 4.2.5 DUPLICATION OF RESEARCH

With respect to the South African context, the survey shows that forty two percent (42%) of the participants thought that ETDs would help minimise

duplication of research, thirty percent (30%) thought they would in some cases and nineteen percent (19%) did not believe ETDs would minimise duplication of research ( see table 4.13 below).

Table 4.13 Results of Q30 of the electronic survey.

|         |               | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|---------------|-----------|---------|---------------|--------------------|
| Valid   | Yes           | 170       | 42.4    | 42.6          | 42.6               |
|         | No            | 76        | 19.0    | 19.0          | 61.7               |
|         | In some cases | 120       | 29.9    | 30.1          | 91.7               |
|         | No opinion    | 33        | 8.2     | 8.3           | 100.0              |
|         | Total         | 399       | 99.5    | 100.0         |                    |
| Missing | No answer     | 2         | .5      |               |                    |
| Total   |               | 401       | 100.0   |               |                    |

Several comments offered by the respondents of the study reveal their thoughts on the question:

“That would depend where it is located / indexed. If not part of for e.g. a widely used/indexed database like Medline, then it would be duplicated somewhere in the world.”

“Duplication might not be a bad thing - different results can be generated.”

“...and it will encourage students to build on what was done before.”

“...the opposite will be the case.”

“...yes unnecessary duplication but at times duplication is needed.”

“The best way of minimising duplication in research is better literature searches. This is primarily dependant on the time invested by the student/researcher.”

“One would hope that before starting research, the person does a thorough literature survey to determine whether or not the work has already been done.

This is however sometimes quite difficult when the research has not been published in easily accessible media. “

“If titles and abstracts are stored in a database, which can be accessed by students, this will make it easier to check for completed or work in progress topics. Otherwise you may have to check every individual University’s website.”

“If the theses are readily accessible, it would be possible to study much of the unpublished work. Many theses contain information that may never be published.”

“Yes it will be easier to access material that may partly overlap, as well as to find relevant research.”

“Sometimes duplication is necessary as it gives different perspectives.”

“I really don’t think duplication of research is a big problem.”

Thus from this qualitative data it appears that that there is not complete agreement on whether duplication of research is necessarily a bad thing in the broader context and furthermore that ETDs which are not indexed in a “interoperable” database may not in any case be found and thus prevent duplication of research.

#### **4.2.6 ACCESSIBILITY AND VISIBILITY**

Eighty seven percent (87%) of the sample agreed that by adding TDs to an online database it would improve the accessibility and use of this research (see table 4.14 page 61).

Table 4.14 Results of Q26 of the electronic survey.

|         |               | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|---------------|-----------|---------|---------------|--------------------|
| Valid   | Yes           | 347       | 86.5    | 87.0          | 87.0               |
|         | No            | 3         | .7      | .8            | 87.7               |
|         | In some cases | 43        | 10.7    | 10.8          | 98.5               |
|         | No opinion    | 6         | 1.5     | 1.5           | 100.0              |
|         | Total         | 399       | 99.5    | 100.0         |                    |
| Missing | No answer     | 2         | .5      |               |                    |
| Total   |               | 401       | 100.0   |               |                    |

Some comments offered by the respondents in respect to Q30 include:

"...depends on where it is indexed."

"Some theses never get published and that information is lost to academics."

"Would improve accessibility, but should be done only after thesis has been accepted by university. "

"Research for post grad theses is often not published. Adding theses to an online database would make them more accessible for a broader readership."

"Most theses are never accessed again - not even the titles."

"In the sciences, most content of theses are also published in journals."

"Access in real time will be possible."

"...should be essential."

"Most definitely!"

"It will eliminate the schlep of having to order through inter-library loan."

"Theses and dissertations more often than not gather dust on library shelves'; making them available electronically contributes to the body of knowledge."

"In the case of my MA it would have been essential - I did an analysis of SABC TV news bulletins recorded on video, which could not be submitted with my thesis as only paper hard bound copies are accepted."

"The interest is more in the publications emanating from the thesis. A thesis does not carry the same weight as does a publication of the same work."

"...it will be easier to find up to date information which can be searched using keywords."

"Online publications circumvent the power of 'established' academic publishers."

"It's the information era and students are familiar with the use of the available internet search engines, thus such online inclusion of theses and dissertations will improve accessibility."

Colleague from Pretoria University states: "Between Dec 2004 and now we have had 1794262 downloads."

"Many relevant dissertations or theses are not available in the library and the abstracts are often the only thing that is available. Authors of dissertations are not always contactable or willing to share electronic copies of work."

"Like electronic journals, theses would be more accessible...the click of a button."

"I prefer peer-reviewed papers originating from theses and dissertations."

"Actually is it cumbersome to get access to other universities' theses and dissertations in the traditional way."

"...must be well catalogued so searches will find relevant information."

An academic from Pretoria University states: "Our students are currently so electronically fluent; it's unthinkable not to let them have access to this type of information."

It is clear from the literature and the survey results that there is a perception that a strong case exists for accessibility benefits of ETDs in the African situation. It is felt that in the first place African scholarship will become better known and shared more widely throughout the world and secondly that ETDs would assist African researchers in gathering data for their own research. However, it is important to distinguish between ETDs that are available free of charge and the publications arising from ETDs that appear only as abstracts as many currently do in existing online journal databases where subscriptions are necessary, in these cases accessibility benefits decrease considerably if research budgets are lacking.

There is also an assumption that simply because content is 'on' a website it will be 'found'. This is an untrue argument as *Mersham (2001:192) argues*. Media habits, including Internet search habits play a vital role, since many people use specific portals to find materials. On the other hand, there is also a possibility that 'Googling' will ultimately become the 'de facto' first choice in the search for research, making some of the existing database portals redundant.

#### **4.2.7 IMPROVEMENT OF RESEARCH PUBLICATION TIMEFRAMES**

Some comments offered by the respondents of this survey included:

"...if downloads are easy and quick, as extensive multimedia may slow things down!"

"It would certainly speed up research and prevent much duplication through ignorance of other research in the field. "

Mersham (2006) argues that while Fox makes important claims for electronic data handling, the reality is sometimes quite different. Many researchers have, for example, 'lost' all or part of their work due to inefficient backup procedures or hardware failures. The claim that 'checking of submissions' is made faster may also not always be true, whether an external examiner examines paper or electronic material, the 'checking' process remains basically the same. In any case, email and electronic editing is widely used in the preparatory work that leads to the physical print version of a dissertation or thesis. Therefore it can be accepted that electronic communications already substantially impact on the development of dissertations and theses whether the final product is an ETD or print version by speeding up the processing of information and interaction between research colleagues.

#### 4.2.8 COST BENEFITS

Figure 4.5 shows that sixty three percent (63%) of the participants of this survey thought that there would be cost benefits in publishing and storing electronic theses and dissertations (ETDs) compared to paper bound theses and dissertations (TDs).

**Q29: Do you think that there would be cost benefits in publishing and storing electronic theses and dissertations (ETDs) compared to paper bound theses and dissertations (TDs)?**

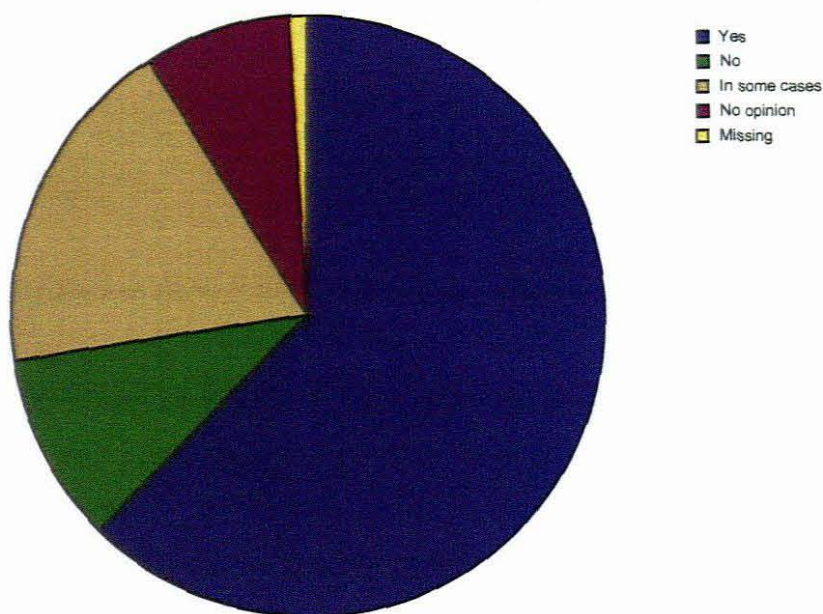


Figure 4.5 Illustrating the results of Q29 of the electronic survey.

Several comments offered by the respondents of this study to Q29:

"It means I have to print and bind these lengthy documents if I wish to read cover to cover or have a copy. Prefer to order from publisher."

"Printing could be more costly than obtaining paper bound copy - cannot read whole dissertation off the screen."

"It will economise on space and time."

"Cost is not the main issue for me - accessibility is. "

"Much cheaper."

"...reduces the paper/manual handling costs."

University of Limpopo colleague states:" According to current practice, the institution would still have to publish the work on paper."

"I speak as a Librarian. As a pro, libraries will have less storage space needs. As a con, students will have greater printing needs."

"The best theses (hard copy) get stolen from the Library - e-theses could not be pinched ."

"...libraries are getting full - new communication medium necessary."

"The man hours going into paper and electronic production is essentially the same."

"Obtaining hard copy from say overseas (via inter library loan facility), for instance, is costly and obviously takes time to arrive."

“From a storage and space usage perspective and a saving on delivery costs between users and institutions.”

“Printing a dissertation of 200+ pages for 5 or more examiners could be costly!”

Colleague from Rhodes University points out: “ETDs require more staff time”

“If it is available electronically no library transfer costs are applicable (these can be very expensive). This makes it easier to access for students who may not have the funding to be able to transfer content relevant to their research.”

“Electronic theses should not substitute provision of hard copies in libraries.”

A postgraduate student states: “...only if the institute purchases the software e.g. PDF creator so there is no additional cost for me.”

“People find it difficult to read from the console hence paperwork will be required that costs money.”

“It would depend on where it is going to be stored and published and the costs involved there.”

Overall it is clear that respondents believe there are cost savings in using ETDs. However the individual comments reveal that in fact there is a range of cost issues involved, for example if training of students is involved. Neither Fox nor any of the respondents mention the real costs to the institution, including running and maintaining the network and servers, which make ETDs possible. These are generally ‘invisible’ to researchers but are nevertheless sometimes quite substantive.

#### **4.2.9 ETD FORMATS AND CONTENT CODING**

Some comments offered by the respondents of this survey included:

Librarian and IT specialist states: "...space, storage of print - is expensive. Problem with digital storage is the format stored now. Will it be accessible in 20 years?"

"Concern of corrupt data etc, loss of data due to viruses, or disks no longer working in 10 years time etc."

A colleague from the University of the Western Cape (UWC) states: "I will not mark a thesis in electronic format. I need a paper copy to mark."

Nico van Aarde from North-West University (NWU) states: "Students must at least submit one hard copy of thesis/dissertation for examining purposes."

A colleague from the University of the Free State (UFS) states: "Printing out a downloaded document makes reading it MUCH easier on the eyes, which undergo enough computer strain!"

"ETDs sound like a fantastic idea whose time has come."

Dr. C.D. Viljoen from the University of the Free State (UFS) writes: "Each generation will differ in its preference with regard to printed material versus e-versions, with the younger generation readily accepting e-versions and the older generation being more resistant in part due to accessibility."

A colleague from the University of KwaZulu-Natal (UKZN) writes: "In the last 5 years since I wrote my thesis there has been a swing from paper-based to digital material. If I were to write another thesis now, my answers to your questions would probably be very different."

A colleague from the University of Cape Town (UCT) writes: "Portability is helpful sometimes. Electronic allows for easy word searching and filtering of relevance."

Professor G. A. Kolawole from the University of Zululand states: "There is always the need for hard copies for future references, since one would have to go on to the Internet for this information. Availability of printable articles/theses online can only supplement access to hard copy materials, it cannot replace them."

"Difficult to read lengthy documents on the screen thus it is easier to take out a copy from the library to read comfortably. Abstracts and synopses are more suitable for electronic formats."

"Electronic storage aids access, but is not viable as a long term storage option. Paper is better. "

"Although I think there will always be a need for paper bound theses to be available too."

"Paper copies will have to stay; universities require that for their collections (so that students who don't have access to Internet can read these). But the electronic media is so powerful and accessible, that it should be used a lot."

Carl Sandrock from Pretoria University writes: "I think you should consider that a fraction of the scientific community uses LaTeX (text formatting system LaTeX is used within the natural and engineering sciences to mark up mathematics and algorithmic graphics in scientific documents) for document preparation and that this does not require Adobe Acrobat. The PDF format is open, so there are many PDF readers and writers. All my documents are available as PDF by virtue of using LaTeX."

Kling and McKim (1999:6) state that while electronic publishing enthusiasts, such as Odlyzko (1996, in Kling and McKim, 1999), consider traditional printed journals "awkward artefacts" that will "likely disappear within ten to twenty years" they view both electronic and paper media as fulfilling a set of useful communicative functions. They believe that each medium provides a package that is convenient for some purposes, and awkward for others, giving

the example of low cost searching and distribution for electronic media compared to ease of reading for paper (Kling & McKim, 1999:6). In terms of Mersham's (1995:55) communication model this is a very relevant point since he argues that the choice of medium is critical to the passage or non-passage of the message content.

#### **4.2.10 PLAGIARISM**

Most (65%) participants of this survey said that they would not have any reservations on publishing their TD online. Thirty seven percent (37.3%) of those who did have reservations stated that plagiarism was the reason for their reservation (see table 4.12 on page 58).

Some comments offered by the respondents of this survey included:

"Strict anti-plagiarism measures need to be put in place."

"ETDs will minimise duplication of research If there are ways of preventing plagiarism ."

#### **4.2.11 TECHNOLOGY AND IMPLEMENTATION**

Several comments from respondents in the survey highlight the lack of technology necessary for the implementation of ETD programmes:

"Many still don't have access to the internet."

"ETDs may not be accessible due to the lack of facilities."

At most South African universities one would expect limited Internet facilities using a quota system for example, to be available to both undergraduate and postgraduate students, while most staff, including library staff would be connected to an Intranet and the Internet. In reality student facilities can often be saturated, or 'bottlenecked' in terms of workstations per student, and management will need to plan for postgraduate computer laboratories dedicated to the ETD lifecycle when implementing an ETD programme.

#### **4.2.12 PRIOR PUBLICATION**

Participants of this survey that said they would have reservations on publishing their TD online, thirty percent (30.7%) stated their reason was that this could be seen as prior publication and ten percent (10.1%) gave patents pending for their reservation (see table 4.12 on page 58).

#### **4.2.13 SOFTWARE APPLICATIONS USED**

Ninety five percent (95%) of the sample were familiar with Adobe Acrobat Reader while (51%) had actually created a PDF file. Only thirty five percent (35%) had created a web page using web authoring programs, with twenty five percent (25%) having published any of this work to the Internet. Software applications being used in the recording of post-graduate work and or research included word processors (32.7%), spreadsheet applications (19.5%), presentation applications (15.7%), database applications (8.3%), Adobe Acrobat reader (14.2%), web authoring applications (3.1%) and other (6.5%).

Software applications being used in the analysis of post-graduate work and or research included word processors (33.8%), spreadsheet applications (30.7%), MS Access (4.8%), SPSS (12.6%) and other (18.1%).

Software applications being used in the presentation of post-graduate work and or research included word processors (29.0%), spreadsheet applications (13.7%), presentation applications (27.3%), database applications (4.7%), Adobe Acrobat reader (11.6%), web authoring applications (2.5%), web - browsers (8.1%) and other (3.2%).

The last question asks participants whether they prefer to print downloaded literature to read. The majority (71%) do, while about a quarter (24%) does in some cases and only five percent (5%) prefer to work with the electronic format (see table 4.15 on page 71).

Table 4.15 shows the response to Q31: Do you prefer to print downloaded literature to read?

|         |               | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|---------------|-----------|---------|---------------|--------------------|
| Valid   | Yes           | 283       | 70.6    | 70.8          | 70.8               |
|         | No            | 20        | 5.0     | 5.0           | 75.8               |
|         | In some cases | 97        | 24.2    | 24.3          | 100.0              |
|         | Total         | 400       | 99.8    | 100.0         |                    |
| Missing | No answer     | 1         | .2      |               |                    |
| Total   |               | 401       | 100.0   |               |                    |

### 4.3 CONCLUSION

I conclude this chapter by noting that the application of frequency statistics to the survey data will limit the assumptions made to the sample surveyed and not to the larger population which would need the application of inferential statistics to the sample data. This oversight is recognised as a limitation of the study.

## **CHAPTER FIVE: SUMMARY, CONCLUSIONS AND RECOMMENDATIONS**

### **5.1 INTRODUCTION**

The study acknowledges that an ETD can take an assortment of formats, at the simplest level it may be a scanned electronic version of a printed thesis or dissertation (TD) that has been converted into Portable document format (PDF). Otherwise it could be a 'born digital' TD produced in a word publisher and converted into PDF, HTML or XML in order to be made available on the World Wide Web. Alternatively it might be created in specialized software applications which are used in specific fields of research like LaTeX in the natural sciences which can also convert to PDF. ETDs may include audio and visual material which will allow researchers to articulate their studies through these media allowing for a 'dual-coded' message which could be perceived as 'richer' to the senses and easier to understand and retain. ETDs may adopt a dynamic information model instead of the traditional linear format found in analogue media such as printed TDs. This will allow for dynamic searching and retrieval of information within the content units of ETDs.

Our preferred methods of communication within the information age promote the advantages of ETDs, emphasising: immediate access, remote access, around-the-clock access, simultaneous use by multiple numbers of researchers and increased use of TDs (Robert Gordon University: 2006). Researchers can undertake full content searches of ETDs on the Web, without having to judge from an abstract whether it is worth requesting a copy of the full text of a hard-copy TD on inter-library loan. (Robert Gordon University: 2006).

Within our academic culture content management systems (CMSs) are becoming increasingly popular to deliver online curriculums where computer facilities are available on campus and in distance learning programmes, which make ETDs the logical progression for postgraduate TD submissions on these systems.

## **5.2 SUMMARY**

### **5.2.1 SAMPLE**

The self-administered survey focused on a sample that had published e-mail addresses and access to and the ability to use technology. Four hundred and one (401) participants took part in this survey from twenty two (22) different tertiary institutions in South Africa. The response rate (8.2%) may be attributed to the fact that an invitation to respond to email or web-based surveys might be perceived as junk mail and mass mailings to published email lists might be perceived as "spam."

We may assume a correlation in some cases between the recipients' interest and awareness in ETDs and their willingness to participate in the survey and the fact that email and web-based surveys may not constitute everyone's preferred method of communication.

### **5.2.2 DEMOGRAPHICS**

The gender of the participants of the study was well balanced with fifty one percent (51%) of the respondents being male and forty eight percent (48%) female. Sixty eight percent (68%) of the participants were White, eighteen percent (18%) Black, six percent (6%) Coloured, four percent (4%) Indian and two percent (2%) Asian. As Leedy (1989:152) warns that in nonprobability sampling there is no way of guaranteeing that each element in the population will be represented in the sample. This study recognises that due to nonprobability convenience sampling, race bias of the respondents seems to have been introduced into the study. However this statistic could also point to a true reflection of the sample of researchers within tertiary institutions in South Africa who would participate in such an e-survey. Forty four percent (44%) of the survey's respondents stated that their home language was English while thirty seven (37%) percent said it was Afrikaans, sixteen percent (16%) was made up of a combination of IsiNdebele (0.7%), IsiXhosa (2.7%), IsiZulu (4.2%), Sepedi (2.7%), Sesotho (2.0%), Setswana (1.2%), SiSwati (0.5%), Tshivenda (1.5%) and Xitsonga (0.5%) which correlates to the race statistics and points to the two main languages of instruction in South African centres of learning being English and Afrikaans.

### **5.2.3 RESEARCH SOURCES**

One of the findings of this study is the preference of seventy four percent (74%) of the participants to online research sources, which they agreed were more efficient than paperbound formats to obtain research literature. The use of peer reviewed sources either in paper or electronic format should form the benchmark or point of reference within any research field. The use of ETDs as research sources is still rather limited with only nine percent (9%) of the participants saying they have downloaded them. Peer-reviewing ETDs on an ongoing basis would effectively give recognition to worthy research and the process could also be used by the author or supervisor to hone any particular study for publication, it could also increase the citation of these studies within mainstream research.

### **5.2.4 PERCEPTIONS OF ETDs**

The concept of ETDs was familiar to about half of those who participated, while forty two percent (42%) had actually browsed digital libraries hosting ETDs only twenty three percent (23%) of the respondents had browsed South African ETD collections. Twenty nine percent (29%) of participants said their tertiary institutions accepted ETDs however forty eight percent (48%) of participants thought that ETDs should be introduced at their tertiary institutions. There was a perception that in some cases by adding multimedia content (graphics, video, sound etc.) to TDs would improve the expression of postgraduate TDs. The majority of the participants (86.5%) agreed that by adding TDs to an online database would improve the accessibility and use of this research source and (63%) believed ETDs could provide a cost effective alternative to printing and binding a large number of paper copies. There was no clear decision on whether ETDs would minimize duplication of research with forty two percent (42%) saying it would, thirty percent saying it would in some case and twenty percent (20%) saying it would not. There seemed to be a perception from some participants that duplicating research to strengthen new concepts or to apply a different angle to the same study was healthy. Most (65%) participants said that they would not have any reservations on publishing their TD online. Those who did have reservations stated that

plagiarism (37.3%), prior publication (30.7%) and patents pending (10.1%) where the reasons for they reservations.

#### **5.2.5 SOFTWARE APPLICATIONS USED IN THE PREPARATION OF TDs**

There was no shortage of client-side software applications being used to record, analyse and present postgraduate work and or research in electronic formats. What is required now is adequate server-side CMSs to host and display these works within accepted web-based interfaces with efficient information models to find and retrieve information.

### **5.3 CONCLUSIONS**

One can conclude that a number of issues arise from the production of ETDs. There seems to be little doubt that South African academics believe that overall ETDs play a positive role in the dissemination of research and that ETDs are poised to make a significant impact on tertiary education in South Africa.

However, there are concerns and differences of opinion about formats, content coding and the question of long-term preservation and transfer of content to future formats. There are also differences of opinion about whether electronic access might encourage greater plagiarism in research. Differences of opinion exist as to whether ETDs tend to reduce the duplication of research or not.

Clearly in South Africa there is a need for the development of programmes to train scholars in the use modern ICT technologies to enable them to create quality multimedia ETDs which can be accessed and used throughout the world. Such programmes should encourage originality and contain ethics training with regard to the issues of plagiarism.

Comprehensive and integrated implementation strategies need to be formally launched at academic institutions to achieve these goals. Additional resources

both human and material will be needed for archive infrastructure, staff and graduate training in the production of ETDs.

There is concern that electronic publication may increasingly be seen by print publishers of academic journals as prior publication and policy will have to be developed in this regard.

All facets of the proposed communication model for ETDs are already functioning within some tertiary institutions of South Africa including Rhodes University, University of Witwatersrand, University of the Free State, University of Pretoria and the University of South Africa (UNISA). For the rest it only seems a matter of time until we follow the worldwide trend of using web-based Content Management Systems (CMSs) for the dissemination of intellectual property. The inclusion of intellectual property rights within good knowledge management policies will also need to be addressed within all tertiary institutions who embrace the idea of ETDs.

The majority of communicators and recipients representing the sample have language and technical skills to encode and decode electronic messages while the minority needs to develop these skills during their academic careers. The electronic format of the messages can be published and accessed using the Internet, which for most participants in the study constituted the most efficient medium of the message. While the scale of access and potential of peer-review and academic recognition on a worldwide scale will only increase the standard and rewards of quality research in South Africa.

For efficient communication of ETDs through portal websites a dynamic living information model needs to be designed around the recipients' needs and preferences. Communicators need to add appropriate metadata to their works to create a broad base to the information which increases recall and recognition of their works while OAI-compliance means all archives created in this way are "interoperable" and can be recalled effortlessly throughout the world wide open archive initiative.

Communicators and library staff need to be taught the advanced technical coding skills necessary to develop ETDs that conform to the requirements of an effective information model. This will include the use of a content management system (CMS) through which the communicator will submit his / her ETD to a portal website. Internal and external navigation techniques and multimedia can be used to enhance the retention and understanding of the message.

An open source CMS like Moodle or features of Google's document sharing could provide an efficient user interface to schedule and document the individual creation of an ETD while opening up publications later for peer review. Tried and tested open source solutions like DSpace and EPrints both have a proven record at successfully creating digital archives for the capturing, storing, indexing, preserving, and distributing a huge variety e-documents (Penman and Milne: 2004).

Thornhill *et al* (2002:13) state that multi-channel communication supporters like Severin (1967) claims that learning is more effective as the number of available cues or stimuli are increased, however when cross-channel cues are irrelevant this will result in inferior communication, this is known as the cue-summation theory. The study recognises that ETDs and their possible multimedia elements which are discussed in this chapter can be communicated through multiple-channels. The vision for researchers and academics should be to complement and add visual richness to static text-and-graphic disseminations as Mayer and Gallini (1990) in Thornhill *et al* (2002:13) conclude that visual cues amplify and explain text, images and facilitate recall of new knowledge.

#### **5.4 RECOMMENDATIONS**

This study recommends a need for all academic institutions in South African to create institutional archives as well as a national hub of ETDs to give our scholars and the country the full benefits associated with accessing local research documentation through electronic networks.

Consideration needs to be given to where the content of the institutions ETD archive will be stored, this study recommends on an in-house server and in a national collection using open source software packages like DSpace and EPrints, which both have a proven record at successfully creating digital libraries for the capturing, storing, indexing, preserving, and distributing ETDs (Penman and Milne: 2004)

Management would have to appoint responsibility for the creation and maintenance of the ETD collection, this study recommends that this should fall under the present library structures. Additional staff maybe needed to train students and staff to create multimedia ETDs in an appropriate form. The text format for the ETD body would currently be Portable Document Format (PDF) as Penman and Milne (2004) state that unlike other electronic file formats such as HTML, PDF preserves the exact layout, font attributes, and formatting of the document from which it was created, ensuring that the electronic version of a document appears just like the original, PDF is also efficient for conversion of paper based documents. XML can be used to define a standard way to add mark-up to ETDs which can allow for common metadata formats to be created, giving institutions worldwide access to powerful searches and the ability to retrieve ETDs (Penman and Milne: 2004).

Capital costs may be incurred for the purchase of a dedicated server and related equipment however most institutions already have the necessary ITC infrastructure and DSpace and E-Prints software are open source and therefore free. The project would require ITC staff time to install the software and to update it periodically. Staff will also be required to provide training for students and staff on an ongoing basis, to assign metadata to ETDs, and to upload the content into the ETD archive. Staff time will constitute the major ongoing cost of an ETD project and therefore management needs to provide budget incentives to the relevant departments within the institution to initially promote the concept of ETDs and to train students to provide quality multimedia content for the archive(Robert Gordon University: 2006).

The study acknowledges that students, academic staff, managerial and administrative staff, and library staff will all have to be involved in an ETD programme and each group will have different priorities or reasons for participating in order for the ETD lifecycle to be completed (Robert Gordon University: 2006).

For students the production of TDs in electronic format provides an opportunity for them to gain new skills and to have their research publicized. The cost effective aspect of ETDs could promote the acceptance of ETDs as students would otherwise be required to provide multiple copies of printed TDs. Students should be assured that a restriction period for an ETD can be agreed upon in case they wish to publish their results in other formats before the full content of the TD is made available on the Web (Robert Gordon University: 2006). The restriction period might eliminate some reservations on ETDs including prior publication and plagiarism following the logic that after publication in a recognized forum it is easier to lay claim to ones copy rights.

Research supervisors should be supported by Library and / or IT staff to provide training and support for postgraduate students who wish to create multimedia ETDs (Robert Gordon University: 2006). However undergraduate curriculums should encompass most of the ITC skills required to produce and submit the simplest versions of 'born digital' ETDs produced in a Word processor, PDF, HTML or XML software application.

Management and administrative staff would be required to introduce and sustain an ETD collection. Generating interest and obtaining the necessary changes to policies and regulations may take a considerable length of time and work (Robert Gordon University: 2006) as many academics would be resistant to change from what they have grown accustomed to.

The Robert Gordon University website (2006) states that for libraries the benefits of ETDs would be that shelf and storage space is saved; electronic access to the content of TDs reduces inter-library loan requests. The usage statistics of the collection would be easier and an improved level of customer

satisfaction could be achieved (Robert Gordon University: 2006). Additional resources, staff and training should be motivated for by recognizing the value and benefits of ETDs to obtain managements support for the creation of an ETD archive and to present a business plan to management to reassure them about the value of the activity and the costs involved. Another key benefit of an ETD archive would be to improve the publicity of research output by the institution's staff and students and, consequently to its general research profile. Attention should be drawn to the number of institutions that are currently in the process of making TDs available electronically and by examining examples of best practice to answer questions and / or concerns about the time involved in the creation of the archive and the ongoing maintenance and development of it (Robert Gordon University: 2006).

## BIBLIOGRAPHY

Africa on the Internet website. Starting Points for Policy Information (2004). [Online]. Available: <http://www.africaaction.org/bp/inetall.html>. Accessed: 23<sup>rd</sup> April 2004.

Bowers, J. (2003), The Ethics of Modern Communication. [Online]. Available: <http://www.jerf.org/writings/communicationEthics/communication2.html>. Accessed: 4<sup>th</sup> August 2004.

Coorough, C. (2001), **Multimedia and the Web - Creating Digital Excitement**. Harcourt College Publishers: London.

Copeland, S. (2004), Electronic Theses: the RGU Project. A Future for UK Theses. [Online]. Available: [http://www2.rgu.ac.uk/library/about/documents/Lon\\_SC.ppt](http://www2.rgu.ac.uk/library/about/documents/Lon_SC.ppt). Accessed: 23<sup>rd</sup> June 2006.

Dalton, J. T. (2001), Electronic Theses & Dissertations: A Survey of Editors & Publishers. [Online]. Available: [http://www.access.wvu.edu/esra/ESRASearch\\_VOpenFile.cfm?serverFilePath=D%3A%2FInternet%2FwwwESRA%2FFiles%5CE87%5CETD%27s%5F%2D%5FA%5FSurvey%5Fof%5FEditors%5Fand%5FPublishers%2Epdf](http://www.access.wvu.edu/esra/ESRASearch_VOpenFile.cfm?serverFilePath=D%3A%2FInternet%2FwwwESRA%2FFiles%5CE87%5CETD%27s%5F%2D%5FA%5FSurvey%5Fof%5FEditors%5Fand%5FPublishers%2Epdf). Accessed: 8<sup>th</sup> December 2003.

Database of African Theses and Dissertations (DATAD), Pilot Project, Final report (2003), [Online]. Available: <http://www.aau.org/datad>. Accessed: 2<sup>nd</sup> December 2003.

Figueres, J. M. (2003), Information and Communication Technologies for African Development: An Assessment of Progress and Challenges Ahead. A Publication of the United Nations ICT Task Force. [Online]. Available: <https://unp.un.org/details.aspx?entry=E04292>. Accessed: 2<sup>nd</sup> June 2003

Fox, E. A., Dobratz S., Guédon J.C., McMillan G., Patel T., Pavani A. and Pockley S. (2001), Guide for Electronic Theses and Dissertations. [Online]. Available: <http://etdguide.org/etdguide.pdf>. Accessed: 19<sup>th</sup> November 2003.

Fox, E. A. (1998), Networked Digital Library of Theses and Dissertations: A Framework for East-West Collaboration. [Online]. Available: <http://citeseer.nj.nec.com/cache/papers/cs/21927/http:zSzzSzwww.theses.orgzSzndltDzSztalkszSzAsia08.pdf/fox97networked.pdf>. Accessed: 17<sup>th</sup> November 2003.

Garfield, E. (1994), The concept of citation indexing: a unique and innovative tool for navigating the research literature. [Online]. Available: <http://scientific.thomson.com/free/essays/citationindexing/concept/>. Accessed: 26<sup>th</sup> April 2006.

Gilbane report, (2002), Content, Computing, and Commerce - Technology & Trends. Vol. 10, No. 1. February, 2002. [Online]. Available: [www.gilbane.com](http://www.gilbane.com). Accessed: 27<sup>th</sup> January 2006.

Global e-Schools community initiative website, (2004). [Online]. Available: <http://www.gesci.org>. Accessed: 2<sup>nd</sup> February 2004.

Harper, D. (2004), The challenge of finding information in (long) e-theses. A Future for UK Theses. [Online]. Available: [http://www2.rgu.ac.uk/library/about/documents/Lon\\_DH.ppt](http://www2.rgu.ac.uk/library/about/documents/Lon_DH.ppt). Accessed: 23<sup>rd</sup> June 2006.

Hill, R. (1998), What sample size is "enough" in internet survey research? Interpersonal Computing and Technology: An Electronic Journal for the 21st Century. ISSN: 1064-4326 July 1998 - Volume 6, Number 3-4. [Online]. Available: <http://www.emoderators.com/ipct-j/1998/n3-4/hill.html>. Accessed: 22nd September 2006.

Kling, R. and McKim, G. (1999), *Scholarly Communication and the Continuum of Electronic Publishing*. [Online]. Available: <http://xxx.lanl.gov/ftp/cs/papers/9903/9903015.pdf>. Accessed: 8<sup>th</sup> April 2004.

Lawrence, S. (2001), *Online or Invisible*. [Online]. Available: <http://www.neci.nec.com/~lawrence/papers/online-nature01/online-nature01.pdf>. Accessed: 9<sup>th</sup> December 2003.

Leedy P.D. 1989. **Practical Research: Planning and Design**, Fourth edition. Macmillan Publishing Company, New York.

Linux home networking website (2006), *Introduction to Networking*. [Online]. Available: [http://www.linuxhomenetworking.com/wiki/index.php/Quick\\_HOWTO:\\_Ch02:\\_Introduction\\_to\\_Networking](http://www.linuxhomenetworking.com/wiki/index.php/Quick_HOWTO:_Ch02:_Introduction_to_Networking). Accessed: 24<sup>th</sup> April 2006.

Matthews J. & Wiggins R. W. (2001), 'Born digital: hypermedia theses'. **Library Journal** 126 no13 Ag 2001. [Online]. Available: <http://www.libraryjournal.com/article/CA149816.html>. Accessed: 7<sup>th</sup> October 2005.

Mersham G.M., Rensburg, R. & Skinner, C. (1995), **Public relations, development and social investment : A southern African perspective**. Pretoria: Van Schaik.

Mersham G.M & Skinner C. (2001), **New Insights into Communication and Public Relations**. Sandown: Heinemann.

Mersham, G.M. (2006), E-mail communication, 29<sup>th</sup> May.

Myers, M. D. (1997), *Qualitative Research in Information Systems*, *MIS Quarterly* (21:2), June 1997, pp. 241-242. *MISQ Discovery*, archival version, June 1997. [Online]. Available:

[http://www.misq.org/discovery/MISQD\\_isworld/](http://www.misq.org/discovery/MISQD_isworld/). Accessed: 17<sup>th</sup> February 2006.

Networked Digital Library of Theses and Dissertations (NDLTD) website, (2005). [Online]. Available: <http://www.ndltd.org>. Accessed: 23<sup>rd</sup> April 2005.

Penman, A. and Milne, R. (2004), I.T. aspects. A Future for UK Theses. [Online]. Available: [http://www2.rgu.ac.uk/library/about/documents/Lon\\_AP\\_RM.ppt](http://www2.rgu.ac.uk/library/about/documents/Lon_AP_RM.ppt). Accessed: 23<sup>rd</sup> June 2006.

Preservation Metadata and the OAIS Information Model (2002), A Metadata Framework to Support the Preservation of Digital Objects. A Report by The OCLC/RLG Working Group on Preservation Metadata. [Online]. Available: <http://www.oclc.org/research/pmwg/> Accessed: 27<sup>th</sup> January 2006.

Rhodes University website, (2005). [Online]. Available: <http://www.ru.ac.za/library/theses/rutheses.html#About>. Accessed: 23<sup>rd</sup> April 2005.

Robert Gordon University Library, Aberdeen, website (2006). [Online]. Available: <http://www2.rgu.ac.uk/library/about.html> . Accessed: 23<sup>rd</sup> June 2006.

Schweig, E. S. *et al* (2001), The Internet: shaking up scientific communication. How a regional mailing list facilitated response to an international event. [Online]. Available: <http://www.nature.com/nature/webmatters/equake/>. Accessed: 17<sup>th</sup> November 2003.

Shannon, D. M. *et al* (2002), Using electronic surveys: advice from survey professionals. *Practical Assessment, Research & Evaluation*, 8(1). [Online]. Available: <http://PAREonline.net/getvn.asp?v=8&n=1>. Accessed: 24<sup>th</sup> March 2006.

Sheth A. P. *et al* (1999), *Processes Driving the Networked Economy*. IEEE Concurrency July–September 1999. [Online]. Available: <http://lsdis.cs.uga.edu/lib/download/SAA99.pdf>. Accessed 26<sup>th</sup> April 2006.

Thornhill, S., Asensio, M. and Young, C. (2002), Video Streaming: a guide for educational development. [Online]. Available: Click and Go Video 2002. Accessed: August 22, 2005.

Ubogu F. N. 2003. From DATAD to ETDs: the Way Forward. [Online]. Available: <http://edoc.hu-berlin.de/conferences/etd2003/ubogu-felix/PDF/index.pdf#search=%22Ubogu%20F.%20N.%202003.%20From%20DATAD%20to%20ETDs%3A%20the%20Way%20Forward%22>. Accessed: 20<sup>th</sup> February 2004.

Ubogu F. N. (2001). 'Spreading the ETD Gospel: A Southern African Perspective'. **The International information and library review**. London Academic Press, Volume / issue: 33 (June) 2-3:249-59.

University of Kentucky's website UK-ETD (2006), Containing Frequently Asked Questions. [Online]. Available: <http://www.uky.edu/ETD/FAQ.html>. Accessed: 17<sup>th</sup> May 2006.

University of Pretoria website, (2005). [Online]. Available: <http://upetd.up.ac.za/about.htm>. Accessed: 23<sup>rd</sup> April 2005

University of Witwatersrand website. 2005. [Online]. Available:  
<http://www.witsetd.wits.ac.za/ETD-db/about.htm>. Accessed: 23<sup>rd</sup> April 2005.

Virginia Tech website. 2006. [Online]. Available:  
<http://etd.vt.edu/howto/publisher.html>. Accessed: 17<sup>th</sup> May 2006.

Virginia Tech website. 2006. [Online]. Available:  
<http://etd.vt.edu/howto/accept.html>. Accessed: 17<sup>th</sup> May 2006.

Wits ETD Pilot Project Final Report. [Online]. Available:  
[http://www.witsetd.wits.ac.za/ETD-db/pilotproject\\_finalreport.doc](http://www.witsetd.wits.ac.za/ETD-db/pilotproject_finalreport.doc) March 2004.  
Accessed: 23<sup>rd</sup> March 2005.

## APPENDICES

### Appendix 1 – Back to chapter 4

Table 4.1 ETD content formats from Virginia Tech website (2006)

|                         |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               |
|-------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Full Text of an ETD     | <p>PDF (.pdf)</p> <p>PDF stands for Portable Document Format. Adobe Systems developed the PDF standard and provides the premiere package for creating and manipulating PDF files, Adobe Acrobat. The process of converting to PDF takes instructions that would ordinarily be sent to a specific printer and prepares them to be viewed or printed on any computer with the free Acrobat Reader installed. The latest version of Acrobat Reader can be downloaded from: <a href="http://www.adobe.com/products/acrobat/readstep.html">http://www.adobe.com/products/acrobat/readstep.html</a></p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |
| Additional Text/Notices | <p>PDF (.pdf)<br/>see above</p> <p>PPT (.ppt)<br/>Microsoft PowerPoint presentations may be included.</p>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |
| Images                  | <p>When deciding what image format to use first weigh its original form to the output. Some formats are better for different data. If you intend that the image will or needs to be printed go for a format that produces a larger file size.</p> <p>General Rule of Thumb:</p> <ul style="list-style-type: none"> <li>• .pdf - good for line drawings with searchable text, e.g. maps.</li> <li>• .jpg - better for photographs</li> <li>• .gif - better for images other than photos, e.g. drawings.</li> <li>• .tiff - for archival images these files are the largest. More info is stored.</li> <li>• .png - created to replace gif format and is acceptable for photos also.</li> </ul> <p>PDF (.pdf)<br/>See above for general information on the PDF format. PDF is best used to store vector-based graphics (i.e. graphics drawn using lines and curves rather than pixels). Vector graphics stored in PDF format will be much smaller, will read more cleanly, and any included text will be searchable. Equations, charts, and diagrams that combine text with vector-graphics are particularly appropriate to store in PDF format.</p> <p>JPEG (.jpg)<br/>The JPEG format is primary used to store photographs. JPEG is a "lossy" format, meaning that some image quality is sacrificed in order to produce much smaller files. Images of higher quality should be stored in TIFF format instead (see below). Non-photographic images such as graphs and charts will be smaller if stored in GIF format instead (see below).</p> <p>CompuServe GIF (.gif)<br/>The GIF format, developed by CompuServe, is best used to store screen-quality images that do not contain many colors. GIF files are typically very small, but cannot reproduce the range of colors necessary to reproduce photographic images (use the JPEG format instead – see above).</p> <p>TIFF (.tif)<br/>The TIFF format is an archival format, meaning that it does not sacrifice image</p> |

|       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
|-------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|       | <p>quality in order to reduce file sizes. TIFF images are excellent for storing detailed, high quality images. However, TIFF files tend to be much larger than either JPEG or GIF images, and cannot be opened using most web browsers without installing and configuring additional viewing software or plugins.</p> <p><b>PNG (.png)</b></p> <p>The PNG format is an open standard developed to replace the Compuserve GIF format. Like GIF files, low-color images stored in PNG format are typically quite small. Unlike the GIF format, the PNG format can also be used to store high-color images, which means it is also suitable for storing photographic content.</p> <p><b>General Suggestions</b></p> <p>For images that are only intended to be viewed on screen, a resolution of 72 or 75 dpi will result in a small file that can be easily downloaded. A resolution of 600 dpi is recommended for images that are intended to be printed.</p>                                                              |
| Video | <p><b>MPEG (.mpg)</b></p> <p>The MPEG format is the oldest and most widely supported format for movies. There are a wide range of viewers available for all platforms. The MPEG format is most commonly used as an output format from UNIX utilities that generate video content.</p> <p><b>QuickTime (.mov,.qt)</b></p> <p>The Quicktime format was originally more of a Macintosh-specific format. These days, support for Quicktime movies is good on both the Macintosh and Windows, but not as good on UNIX.</p> <p><b>Audio Video Interleaved (.avi)</b></p> <p>The AVI format is more of a Windows-specific format, and is not as well supported on other platforms.</p> <p><b>Suggestions Regarding Video Content</b></p> <p>Video is one of the most resource-intensive types of multimedia. Unfortunately, video content that is of even half broadcast quality is often too large to download from home. Consider including lower quality versions of video content in addition to high quality originals.</p> |
| Audio | <p><b>WAV (.wav)</b></p> <p>The default standard for Windows sound files, is also supported for most other platforms.</p> <p><b>AIFF (.aif)</b></p> <p>The AIFF format is a Macintosh-specific equivalent of the WAV format. It is not as well supported on all platforms as the WAV format.</p> <p><b>MPEG-3 (.mp3)</b></p> <p>MPEG-3 (or MP3) format eliminates sound data which is not as strongly perceived by the human ear and brain, and as such creates files of reasonable quality that are as much as 10 times smaller than the raw data itself. MP3 files are good for storing long passages of sound content where high quality is not required.</p> <p><b>Suggestions Regarding Sound</b></p> <p>The quality used to store sound in electronic format reflect the quality of the original</p>                                                                                                                                                                                                                |

|               |                                                                                                                                                                                                                                                                                                                                                                                                                                     |
|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|               | recording source. There is very little reason to store low fidelity recordings of speech content in a very high-quality format, as the added file size would not result in any increase in quality. Conversely, high-fidelity recordings should be stored at high-quality.                                                                                                                                                          |
| Other Formats | If you have content that has been created in a proprietary format, it is recommended that you include a copy of the content in both the proprietary format and in a more common format as well. If you have multimedia content that is too large to be downloaded via the web, it is recommended that you include a copy of the content stored at the original quality, and a copy stored at lower quality (and hence smaller size) |

Appendix 2: URL to study's website and online resources:

<http://www.lis.uzulu.ac.za/etd>