e-RESEARCH:
AN IMPLEMENTATION FRAMEWORK FOR SOUTH AFRICAN ORGANISATIONS

A Research Report presented to the
Graduate School of Business Leadership
University of South Africa

In partial fulfilment of the requirements for the
MASTERS DEGREE IN BUSINESS ADMINISTRATION,
UNIVERSITY OF SOUTH AFRICA

By
S FERNIHOUGH
MAY 2011
DECLARATION

I, Shelly Fernihough, declare that this Research Report is my own work, except as indicated in the references and acknowledgements. It is submitted in partial fulfilment of the requirements for the Masters Degree in Business Administration to the Graduate School of Business Leadership (University of South Africa). It has not been submitted before for any other degree at this or any other university. This report was prepared under the guidance of Lorrayne Duweke at the Graduate School of Business Leadership (University of South Africa).

Shelly Fernihough
02 May 2011
ACKNOWLEDGEMENTS

This work is dedicated to my fiancée Graham, family and friends, for their constant support and encouragement during my studies. Thank you for believing in me.

I would like to extend my sincere gratitude to the following individuals:

• Dr Martie van Deventer for her kindness in mentoring, guiding and supporting me through this study, without you this study would not have been possible.

• Lorrayne Duweke, my supervisor, for her guidance, encouragement and direction on this research project.

• To the focus group attendees and interviewees, for giving up their time to share their knowledge and expertise with me.
ABSTRACT

The South African Department of Science and Technology’s corporate strategy 2010-2013 (DST, 2010) places an emphasis on the need for South Africa to become a “knowledge-based economy”. The strategy highlighted the need for integrated infrastructure planning in order to move the country from being a resource-based economy to one that is based on knowledge (DST, 2010). The South African strategy is not entirely unique. To achieve this end many countries have simultaneously identified and are implementing an e-Research strategy.

e-Research is a new way of doing research, collaborating globally and nationally while making use of ICT infrastructure to do research. This research project set out to understand the dynamics involved in e-Research, and gain a better understanding of what the various components are that make up the e-Research paradigm, while at the same time looking at how these various components are funded. To this end the research focused on studying countries with established e-Research initiatives.

In order to fully understand the components of e-Research, a comprehensive literature review was undertaken. An existing framework of e-Research was elaborated on and extended. To calibrate this framework, a focus group discussion was conducted, and to further supplement the focus group’s findings, semi-structured interviews were conducted.

The research results led to a better understanding of e-Research holistically and allowed for the establishment of a comprehensive framework for e-Research was established. The results further revealed aspects that can not easily be depicted on the framework, such as the need for an e-Research leader, to drive the required change. The results also indicated that a lack of appropriate funding
could inhibit the implementation of e-Research, and that much work would be required on developing funding policies and models to ensure the successful implementation of e-Research. The research further indicated that high levels of collaboration between all stakeholders would be required, including amongst others, government, research institutions, and universities.

This framework is intended to assist the South African Government and selected institutions in driving the e-Research paradigm forward at an accelerated pace. It can further assist managers in implementing an e-Research strategy within their own organisations. Lastly the framework could assist in the development of strategy specific to and in support of e-Research.
# TABLE OF CONTENTS

**CHAPTER 1 : DEFINITION OF THE RESEARCH PROBLEM .......... 1**

1.1 Introduction and Background ........................................ 1
1.2 Context of the Problem ............................................. 2
1.3 Problem Review ...................................................... 4
1.4 Research Problem .................................................... 6
1.5 Research Objectives .................................................. 6
1.6 Delineation of the Study ............................................. 7
1.7 Importance of the Study ............................................. 7
1.8 Limitations of the study ............................................ 8
1.9 Outline of Chapters .................................................. 9
1.10 Summary ............................................................... 10

**CHAPTER 2 : THEORETICAL CONSIDERATIONS ..................... 11**

2.1 The Context of Change ............................................... 11
2.2 e-Research Strategy Implementation ............................... 15
   2.2.1 Structure ......................................................... 16
   2.2.2 Culture .......................................................... 16
   2.2.3 Systems .......................................................... 17
       2.2.3.1 Information Services ........................................ 17
       2.2.3.2 Multi-disciplinary Knowledge Management ............... 20
       2.2.3.3 Cyberinfrastructure Services .............................. 22
2.3 Summary ............................................................... 24

**CHAPTER 3 : LITERATURE REVIEW .................................. 25**

3.1 e-Research Infrastructure & Systems .............................. 25
   3.1.1 Grid Computing ................................................. 26
   3.1.2 Networks ........................................................ 30
   3.1.3 High Performance Computing ................................... 36
   3.1.4 Middleware ...................................................... 38
3.2 Managing Data, Information and Knowledge ....................... 41
   3.2.1 Digital Repositories .......................................... 44
   3.2.2 Access to Information .......................................... 46
   3.2.3 Digital Curation & Preservation ............................. 49
3.3 Tools and Applications ............................................. 51
  3.3.1 Communication and Collaboration .............................. 52
  3.3.2 Visualisation ..................................................... 53
  3.3.3 Project Specific Tools and Applications ......................... 55
  3.3.4 Access, Authentication and Authorisation ....................... 56
3.4 Trends in e-Learning and Digital Scholarship .................. 57
3.5 Funding Structures & Oversight of e-Research Activities ..... 59
  3.5.1 Overview of Global e-Research Funding ......................... 59
  3.5.2 Australian Funding Structure of e-Research Activities ...... 61
  3.5.3 UK Funding Structure of e-Research Activities ................ 62
  3.5.4 South African Funding Structure for e-Research Activities 64
  3.5.5 Coordination and Oversight ....................................... 65
3.6 Culture ........................................................................ 67
  3.6.1 Leadership .......................................................... 67
  3.6.2 Technical Expertise and Skills Development ................. 69
3.7 Summary .................................................................... 71

CHAPTER 4 : RESEARCH METHODOLOGY ................................. 73

  4.1 Introduction ........................................................... 73
  4.2 Research Design Principles ......................................... 73
  4.3 Research Approach and Method ..................................... 74
  4.4 Data Collection Methods ............................................ 77
  4.5 Population and Sample ............................................... 78
  4.6 Data Analysis methods ............................................... 79
  4.7 Validity and Reliability ............................................. 79
  4.8 Summary .................................................................... 80

CHAPTER 5 : RESEARCH RESULTS & DISCUSSION ...................... 81

  5.1 Introduction ........................................................... 81
  5.2 Findings for Research Objective 1: UK and Australian Components of e-Research & Funding ............................. 82
    5.2.1 Components of e-Research ......................................... 82
    5.2.2 Sources of Funding for e-Research ............................. 83
  5.3 Findings for Research Objective 2: South African Components of e-Research & Funding ............................. 85
5.4 Findings for Research Objective 3 & 4: A calibrated framework of components for e-Research ................................. 86
  5.4.1 Components of e-Research ............................................. 87
    5.4.1.1 e-Research Infrastructure or “Cyberinfrastructure” Components ................................................................. 88
    5.4.1.2 Middleware & Services ............................................. 89
  5.4.2 e-Research Applications ............................................... 89
    5.4.2.1 Products and Services ............................................. 90
    5.4.2.2 Access & Mobile/Remote Connectivity Components ...... 91
    5.4.2.3 Quality Assurance, Training & Skills Development Components ................................................................. 92
    5.4.2.4 Users, Researchers and the Virtual Research Environment ................................................................. 94
    5.4.2.5 Strategic Oversight, Leadership & Co-ordination Components ................................................................. 95
    5.4.2.6 Collaboration Component ............................................. 96
  5.4.3 Funding of e-Research Components .................................... 97
  5.5 An e-Research Implementation Framework for South African Organisations ................................................................. 99
  5.6 Summary ............................................................................. 110

CHAPTER 6 CONCLUSIONS AND RECOMMENDATIONS .............. 111
  6.1 Introduction ...................................................................... 111
  6.2 Conclusions ...................................................................... 111
  6.3 Recommendations ................................................................ 115
  6.4 Closing Remarks .................................................................. 118

REFERENCES .............................................................................. 119

APPENDICES ............................................................................. 131

Appendix 1: .............................................................................. 131

Appendix 2: .............................................................................. 139
# TABLE OF FIGURES

Figure 1: The Change Kaleidoscope (taken from Mabey, 2007, based on Balogun et al., 1998) ................................................................. 13  
Figure 2: The Strategy Implementation Process (Boojihawon, 2007) .... 15  
Figure 3: Components of e-Research (Page-Shipp, et al., 2005) ....... 18  
Figure 4: Model of a Knowledge Economy (Mbananga, 2007) ........... 21  
Figure 5: An Integrated Cyberinfrastructure Services Model (Adapted from Atkins, 2005) ................................................................. 22  
Figure 6: Conceptual Aren Backbone (Sargent, 2002) ...................... 31  
Figure 7: SuperJanet4 Backbone (Hey and Trefethen, 2003) .......... 32  
Figure 8: UK Regional and National Network Architecture (Sargent, 2002) ....................................................................................... 33  
Figure 9: South African National Research and Education Network (Wright, 2010) ........................................................................ 35  
Figure 10: Basic ICT Infrastructure for e-Research ........................... 38  
Figure 11: Possible Middleware and Service Components of the e-Research Framework ....................................................................... 51  
Figure 12: Tools and Application Components for e-Research .......... 58  
Figure 13: Investments by Nations in e-Research Initiatives (Alberta’s Cyber Infrastructure Taskforce, 2006) ........................................ 60  
Figure 14: Structure and Funding for UK e-Science Programme (Hey and Trefethen, 2002) ................................................................. 63  
Figure 15: e-Research Governance and Skills Development .............. 71  
Figure 16: e-Research Infrastructure or “cyberinfrastructure” ........... 88  
Figure 17: e-Research Middleware and Services ............................. 89  
Figure 18: e-Research Applications ................................................... 90  
Figure 19: e-Research Products and Services .................................... 91  
Figure 20: Access Modules for e-Research ....................................... 92  
Figure 21: User Training, Skills Development & Quality Assurance Components ......................................................................................... 93  
Figure 22: Users and the Virtual Research Environment for e-Research ......................................................................................... 94  
Figure 23: Committee and Management Team for e-Research .......... 96
Figure 24: Collaboration in e-Research Activities .......................... 97
Figure 25: Funding of e-Research Components ............................ 99
Figure 26: An e-Research Implementation Framework for South African Organisations .......................................................... 101
Figure 27: Preliminary e-Research Implementation Framework ...... 133
GLOSSARY

Curation | Curation encompasses many of the preservation and archiving activities, especially in digital environments. Curation is broadly interpreted as being about maintaining and adding value to a trusted body of digital information for current and future use (Borgman, 2007).

Cyber infrastructure | Refers to the distributed computer, information and communication technologies that provide the platform on which to build the new types of scientific and engineering knowledge environments which will enable research to be conducted in new ways and with increased efficiency (DSTC, 2004).

Digital Scholarship | A ‘networked’, scholarly or academic environment with pervasive integration of digital technologies into everyday learning (Mutula, 2009).

e-Learning | Any means of enhancing learning via the use of information and communication technologies (Kellner, 2004 in Borgman, 2007).

e-Research | The more general version of e-Science and includes non-scientific research, such as humanities and social sciences, and is also characterised by the need to use distributed computing resources for collaboration and sharing of knowledge (DSTC, 2004).

e-Science | The term that is given to large scale science that is increasingly being carried out through distributed, global collaborations, enabled by the Internet and related technologies (DSTC, 2004).

Grid technologies | Technologies that play an important role in the development of e-Science and e-Research. In much the same way that consumers and businesses are able to access their electricity supply, Grids enable
researchers and research institutions networked access to distributed data repositories, specialised scientific equipment, knowledge services, and computing power in an ‘on demand’ way. They enable flexible, secure resource sharing and coordinated problem solving amongst dynamic collections of individuals, institutions and resources – often referred to as Virtual Organizations (DSTC, 2004).
CHAPTER 1: DEFINITION OF THE RESEARCH PROBLEM

1.1 Introduction and Background

The South African Department of Science and Technology’s corporate strategy 2010-2013 (DST, 2010) places an emphasis on the need for South Africa to become a “knowledge-based economy”. The strategy further highlighted the need for integrated infrastructure planning in order to move the country from being a resource-based economy to one that is based on knowledge (DST, 2010). Minister Pandor, in the foreword to the strategy, particularly mentioned that Research & Development promotes the growth of knowledge-intensive activity in the South African economy. She went on to declare that her department’s aim is to, through research, improve the competitiveness of the South African economy significantly.

This shift of strategy, to be more focused on knowledge development, is not unique. Countries such as the United Kingdom (UK) and Australia followed a strategy: of using research, knowledge and information technology to add value to its resource-based economy (Government of the Republic of South Africa, 2002: 35). A variety of programmes were launched by the aforementioned countries to identify and then address these needs.

Simultaneously, huge changes have taken place in information and communication technology (ICT). These changes have brought about research that was not previously possible and has led to the adoption of a new paradigm known as e-Research. According to Hussey and Hussey (1997: 47), a paradigm refers to the progress of scientific practise based on people’s philosophies and assumptions about the world and the nature of knowledge; in this context, about how research can be conducted.
Different countries use different terminologies for this new paradigm such as e-Science, cyberinfrastructure and e-infrastructure. Meyer and Schroeder (2009) indicate that e-Research has become an all-encompassing or “umbrella” term used to mean the use of advanced computing tools and high powered networks for collaboration and sharing resources in scientific or academic research for the collaborative production of knowledge” (Borgman, 2007; Dutton and Meyer, 2009; Schroeder, 2007).

The e-Research initiative aims to revolutionise research and allow researchers to share and coordinate human and technical resources (Perrott and Harmer, 2008). e-Research is a new way of doing research, collaborating globally and nationally while making use of ICT infrastructure to do research. The Australian e-Research Coordinating Committee (Australian Government, 2006) declared that e-Research has the potential to increase the efficiency and effectiveness of research endeavours across all disciplines. Greater interactivity between researchers and an increased ability to access research outputs will benefit industry, governments and the Australian community as a whole. It stands to reason that the South African Government’s drive to improve research infrastructure and support would address all aspects identified as prerequisites for successful e-Research.

1.2 Context of the Problem

Once addressed at national level, research institutes such as the Council for Scientific and Industrial Research (CSIR) and the Human Sciences Research Council (HSRC), or academic institutions such as the University of Pretoria (UP), would in turn need to create a strategy in support of the wider national initiative. In order to develop and implement such a strategy, the necessary systems, structure and culture would need to be in place to ensure successful implementation. However, as e-Research requires collaboration between many
stakeholders for successful implementation, merely implementing it in one organisation would not truly add value to the knowledge economy.

In searching for previous South African investigations into e-Research, as a holistic concept, only one early endeavour was found.

Research done by the South African Research Information Service (SARIS) project team, found that to participate effectively and to advance e-Research, will require that South Africa as a whole and not only individual institutions will need to invest in technically complex and often costly infrastructural support services (Page-Shipp, et al., 2005). A joint approach towards e-Research by the major stakeholders is desirable, and there is great benefit to be obtained by sharing development and operating costs (Page-Shipp, et al., 2005). The SARIS project team further identified the various components making up the “family of e-Research”, which includes e-Science (the generation of scientific data), the ICT infrastructure, tools and applications, digital curation and preservation, as well as access to e-Information. Their model, in which each of these components is described in more detail, is provided in section 2.2.3.1 of this report.

According to Page-Shipp, et al., (2005), the international community has been grappling with the requirements for e-Research support for some time. As such, South Africa should perhaps look to other countries with more established e-Research programs for guidance on the best practices for e-Research infrastructure and support. Funding e-Research activities, unless coordinated, can be a costly and fruitless exercise (Page-Shipp, et al., 2005). The SARIS team then developed a framework of e-Research which could serve as an initial attempt to design a national framework.

As no further proof could be found that a coherent framework for South African e-Research is openly accessible, it sparked the question whether South African research organisations (such as the CSIR and the HSRC), and academic institutions (such as UP), as key
stakeholders and implementation facilities for national research endeavours, would be willing to assist in developing a national framework into which the development of their own e-Research endeavours could be tied in as a first step into the investigation, it would be necessary to focus on obtaining a detailed understanding of the components of e-Research. The components described by Page-Shipp, et al., (2005) were found to be very basic, and possibly not sufficient to serve as a holistic framework that could guide stakeholders in the national system of innovation, such as the CSIR, HSRC and UP in their strategy development.

1.3 Problem Review

Many countries are funding e-Research in order to build a knowledge economy in their countries. However to build a knowledge economy calls for the integration of many disciplines including not only in specific subject disciplines (such as physics or mechanical engineering) but also information management and knowledge management (Mbananga, 2007). Multidisciplinary, in addition, calls for new ways of thinking about research infrastructure.

Twenty Five years ago, the ICT infrastructure was well suited to the science/research culture of that era (Abbott, 2009). Volumes of data were relatively small; and IT systems were relatively expensive and limited to being accessible only to experts (Abbott, 2009). In the knowledge era, however, many researchers are affected by what is commonly referred to as the data deluge. The data deluge, according to Hey and Trefethen (2003), refers to the spectacular growth of data over the last century.

The volume of scientific data being generated by highly instrumented research projects (linear accelerators, sensor networks, satellites etc.) is so great, that it can only be captured and managed using
information technology (Hey and Trefethen, 2003). The amount of data produced far exceeds the capabilities of manual techniques for data management, and thus the need for control of this data is another essential driver of e-Research (Lord and Macdonald, 2003). This ‘data deluge’ is one of the main drivers of e-Research (Hey and Trefethen, 2003).

Advances in technology are further drivers of e-Research. Distributed working environments, supported by technology, are emerging to break down the barriers previously imposed by distance and time (NSF, 2003). These new research environments are linking together research teams, digital data and information libraries, high performance computational services and scientific instruments (NSF, 2003).

e-Research is continuously evolving, and what appears to have initially started as merely the ICT support to allow researchers to collaborate, has been extended to incorporate the need for open access to information, the need for support of the complete scientific lifecycle, and the need for knowledge production, management and movement of information. e-Research provides opportunities to develop whole new areas of valuable research and to see existing research in new ways (O'Brian, 2005).

Countries, such as the UK and Australia, have established e-Research programs to meet the changing needs of researchers, by using advanced technologies. These countries have made great strides towards implementing e-Research nationally and institutionally. Their programs are constantly evolving, and the lessons learnt from these programs could possibly assist developing countries in establishing their e-Research programs more efficiently and effectively.

Before implementing e-Research in South Africa, an evaluation phase should first take place. Looking at countries with established programs in order to evaluate their programs can provide insightful information
for South Africa, as a leader in Africa. An evaluation of established programs, while also taking into consideration African circumstances, can provide useful information pertaining to the implementation requirements for South Africa, in order to place South Africa on the forefront of e-Research in Africa.

An evaluation of established programs could further provide South Africa with insight into the most cost effective way to implement e-Research, as well as looking at the lessons learnt from implementing e-Research abroad. As such, South Africa should look to other countries with more established e-Research programs for guidance on the best practices for e-Research infrastructure and support.

1.4 Research Problem

e-Research: An Implementation Framework for South African Organisations

1.5 Research Objectives

➢ To investigate and identify the components of the e-Research and to determine sources of funding for these components, as identified in the United Kingdom (UK) e-Science, and the Australian e-Research programs.
➢ To investigate and identify the components of the e-Research, as implemented in South Africa.
➢ To elaborate on and extend the framework of components of e-Research for South Africa and make recommendations for those components that should be funded centrally by government, and those that should be funded by the individual institutions.
➢ To calibrate the framework and suggested funding division (government vs. organisation level) with at least one large research institution.
1.6 Delineation of the Study

This study looks at e-Research holistically from a National perspective. Implementing e-Research in one organisation is not sufficient towards building a knowledge economy and as such e-Research will affect many organisations, but needs to be addressed at an organisational as well as national level.

For these reasons the study focuses on the National perspective but also focuses on issues which organisations who implemented an e-Research strategy need to consider. The findings and recommendations of this study are aimed at government, universities, research institutes libraries and information and communication (ICT) departments to assist them in their e-Research strategy development and implementation. The framework can then be adapted to meet specific institutions needs.

1.7 Importance of the Study

This study will be a significant endeavour in understanding the infrastructure required to fully participate in, and benefit from, e-Research, as a key requirement for building a knowledge economy for the country. The study will also look into intangible aspects, such as culture, which were possibly previously not considered.

The study aims to identify gaps in the framework of components of e-Research, as identified by the SARIS project team. The study further endeavours to show how, having a complete understanding of the various components that make up the e-Research paradigm, can assist universities, government, libraries and research institutes in drafting their e-Research implementation strategies and determining what funding is required and which components are the responsibility of each of the stakeholders.
1.8 Limitations of the study

Desk research will mainly be conducted. It will not be possible to have direct access to any of the architects of the UK or Australian initiatives. It is therefore possible that important issues could be underestimated or not be identified.

Another limitation is that the comparative study will focus on developed countries (as indicated in section 1.3), where larger scale funding is available for implementing such initiatives. In developing countries such as South Africa, the availability of funding is limited and priorities are different.

The study will focus on programs and initiatives that have been implemented by the countries being studied; it will not look into the success or failure of those projects, and as such, will only provide an overview, as opposed to indicating whether the implementation of a specific component is essential.

Time limitations will result in a very limited calibration of the framework and recommendations. Ideally, it would be advisable to test the framework with more institutions.
1.9 Outline of Chapters

This dissertation consists of six chapters.

Chapter 1 provides a general introduction and outlines the problem and background related to the study. This chapter also outlines the research objectives around which the study is structured.

Chapter 2 explores and develops a deeper understanding of the research problem, through reflection and exploitation of appropriate business models and theory.

Chapter 3 provides a comprehensive literature review of the major areas addressed in the investigation. The chapter highlights those aspects that authors indicate as being components of the e-Research paradigm.

Chapter 4 explains the research design and methodology used for the study. Data collection and analysis methods for this study are explained in this chapter.

Chapter 5 sets out the overall results of the study. The results from the focus group conducted and interviews are indicated in this section. This chapter further outlines an implementation framework for e-Research in South Africa.

Chapter 6 discusses the conclusions of the study and well as making recommendations for taking the framework forward in South Africa, while also proposing recommendations for future research.
1.10 Summary

This chapter highlighted the background as to how e-Research came about and explained that this was largely due to the changing needs of researchers, globalisation, increasing volumes of data and improving technologies. The chapter further indicated that countries such as the UK and Australia have been employing this new paradigm – e-Research, as a way to attempt to contribute towards building knowledge economies for those countries. The chapter further emphasized the objectives and the importance of the study, while also providing the outline of the chapters for this report.

The following chapter applies various business models, in order to gain a deeper understanding of the various aspects that affect the implementation of e-Research.
CHAPTER 2 : THEORETICAL CONSIDERATIONS

The potential for advanced ICT to enhance research outcomes and to facilitate new forms of research and research collaborations touches every field of research endeavours, from natural sciences to arts (Sargent, 2007). The needs of these various research activities, the differing perspectives of these needs, and the balances between physical/digital and past/future contexts, are immensely diverse and often divergent (Sargent, 2007). However, the fundamentals of access, collaboration and exchange, are consistent across this diversity.

To achieve effective outcomes for investments in e-Research, and to build platforms for not only current, but also future needs, requires a clear strategic focus (Sargent, 2007). This chapter reflects on existing theories and models, especially how they relate to the e-Research initiative, as well as those aspects that affect the successful implementation of an e-Research strategy.

2.1 The Context of Change

The external forces on organisations require them to change in order to remain competitive (Sturges, 2007). The same can be said for countries. In order to remain competitive globally, countries also need to change. The main external driving forces which cause countries to need to adapt and change are mainly globalisation and changing technology (Sturges, 2007).

Sturges (2007) indicated that while technology is forcing organisations and countries to change, many change initiatives are closely linked to the introduction of new technology. Beyond the introduction of new technologies in business, another approach being used to bring about change is that of knowledge management, in order to achieve competitive advantage. Knowledge management makes use of technologies to harness skills and competencies. e-Research can be
seen as a combination of these change initiatives; both the introduction of new technologies and implementation of knowledge management.

For organisations or countries to implement change, Mabey (2007) indicated that the change initiative should be managed. Mabey (2007:110) notes that the management of change entails a “...series of informed choices...”, whereby the most appropriate approach is selected, and thereafter the tactics to be employed are identified. These areas are illustrated in Figure 1. The eight contextual factors depicted in the outer circle influence the implementation of change and are considered especially important in the early diagnostic phases of change (Mabey, 2007).

Decisions regarding scope, time, power and readiness are considered contextual in nature, while those relating to capacity, capability, uniformity and preservation are seen to rely more on the change leader’s judgment. However, as Mabey (2007) points out, the outcome of transformational change is not necessarily predictable, especially where it entails change at several levels.
The change kaleidoscope refers to scope as being one of the dimensions that requires consideration before bringing about change. When considering scope in relation to change, Mabey (2007) indicated that a judgement must be made on:

- How big a change is called for;
- Whether the change affects discrete working teams or groups;
- Whether the change addresses issues at an inter-group level or at an organisational level.

In answering these questions, relating to scope, implementing e-Research calls for a change which affects many institutions, especially research and academic institutions. The change is wide-spread and many stakeholders will be affected.

In terms of time, the time taken to fully implement e-Research will be dependant on the funding available to implement the necessary, and often costly, infrastructure required. However, the longer it takes to
implement, the more the country or organisation may slip behind its global competitors.

The change kaleidoscope also highlights the need to look at power. Mabey (2007) indicated that power refers to the power the change leader or agent has in order to impose change and how much room the change leader has to manoeuvre. Mabey (2007) also highlights the need for a change leader. Mabey (2007) further indicated that organisations and their people need to be aware of the need for change, as well as being committed to the change.

In terms of capacity and capabilities, it can be highlighted that e-Research brings about new patterns of work, a new way of doing research, and as such, researchers may not have the necessary skills to make use of e-Research capabilities. This will call for new skills development.

From applying the change kaleidoscope, it has been highlighted that a change leader may be necessary, that funding can impact the time taken to implement the change and that their needs to be investment in skills development, to enable researchers to effectively use e-Research.

According to Sturges (2007), based on a classification by Thurley and Wirdenius (1973) in the e-Research context, a strategy for change to best bring about the e-Research initiative may be a participative strategy. The underlying assumption of this strategy, according to Sturges (2007), is that people will not be committed to the change unless they are involved. As e-Research affects a number of stakeholders and the underlying concept of e-Research is collaboration, a participative strategy may be the best way to bring about change.
In order to bring about change on such a large scale one needs to devise a strategy to start implementing the change strategy.

**2.2 e-Research Strategy Implementation**

Before a strategy can be implemented, such as an e-Research strategy Boojihawon (2007) indicated that the necessary structure, systems and culture should be in place (See Figure 2). Van Deventer and Pienaar (2008) indicate that the recommendations from the SARIS report implied that individual research institutions should implement e-Research strategies, collaborate with other institutions and lobby government to support open access initiatives.

![Figure 2: The Strategy Implementation Process (Boojihawon, 2007)](image)

The following sections apply this model to unravel the complexity relating the implementation of an e-Research strategy. Sections 2.2.1 to 2.2.3 below, describe each of the aspects in more detail.
2.2.1 Structure
As was highlighted by the change kaleidoscope (Figure 1 in Section 2.1), leadership may be necessary to effectively bring about change, and while leadership is not often seen to form part of structure, but rather culture, it talks to the need in a national context for a department/institution/body which will lead the change process and will coordinate e-Research activities to eliminate duplication of activities and ensure the cost-effective implementation of e-Research.

Furthermore, structure in the e-Research context can also relate to funding by indicating how e-Research will be funded and who the funding bodies will be. This can also relate to a funding structure for e-Research activities.

The basis of e-Research, according to many authors, is that it allows for collaboration among institutions. This means that for specific research projects, people will be structured in different ways, dependant of the project being undertaken. According to the e-Infrastructure Reflection Group (2010), the organisational structures will need to be service-orientated across organisational boundaries, as well as addressing the integration of local, global and outsourced resources. This may call for dynamic structures across organisations to be established for short periods of time to fulfil a specific purpose.

2.2.2 Culture
It was highlighted in the change kaleidoscope (Figure 1 in Section 2.1) that there would possibly be a need for skills development. However, this change calls for a different way of doing work and may call for a culture change in organisations. Many researchers may not be used to applying advanced ICT technologies to perform their work and bringing about a change as large as e-Research, could have a very big impact on the researchers. Nationally and within organisations, e-Research calls for the need of a possible culture change.
2.2.3 **Systems**

Many authors place an emphasis on the infrastructure required to bring about e-Research. In order to understand e-Research in its entirety, however, means that aspects beyond just infrastructure should be considered. Three disconnected models were considered to gain insight into what could possibly serve as a holistic e-Research systems model. Each of these is described in more detail below.

2.2.3.1 **Information Services**

Page-Shipp, *et al.*, (2005) indicated that the e-Research paradigm is composite of the following trends:

- The ability to transfer large volumes of data;
- To share computation capacity between remotely situated researchers in order to achieve *'faster, different, better research'*;
- The need to make better use of expensively created databases by *'the active management and appraisal of data over the life cycle of scholarly and scientific interest'* is the basis of a new field of endeavour called digital curation;
- Scholarly discourse now takes place on the dual playing field of commercial publication and open access.

The SARIS project team indicated a basic framework for the components of e-Research (Page-Shipp, *et al.*, 2005).
**Components of eResearch**

<table>
<thead>
<tr>
<th>Data Transfer &amp; Computation</th>
<th>Tools &amp; application</th>
<th>Primary Data Sharing</th>
<th>Digital Curation &amp; Preservation</th>
<th>Access to eInformation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>eScience</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>Access to eInformation</strong></td>
</tr>
<tr>
<td><strong>Digital Curation &amp; Preservation</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>Commercial Publishing</strong></td>
</tr>
<tr>
<td><strong>Open Access Publishing</strong></td>
<td></td>
<td></td>
<td></td>
<td><strong>Open Access Publishing</strong></td>
</tr>
</tbody>
</table>

**By definition this is**

- **Science employing transfer and sharing of large volumes of data**
- **Software that allows manipulation, modelling and analysis of data**
- **Making research data available to other researchers**
- **Active management of databases including promotion of effective and widespread use of the datasets for their scientific & scholarly useful life**
- **Contribution to & use of published resources requiring payment by readers**
- **Contribution to & use of published resources where content is regarded as ‘free’**

**Which requires**

- **Access to remotely held large datasets & high performance computing via affordable high bandwidth**
- **Access to models, source code and open standards**
- **Accessible repositories & quick reference**
- **Preservation & curation repositories & access mechanisms, archival skills & infrastructure**
- **Affordable licences for researcher access & discovery mechanisms**
- **Serviceable infrastructure for publication and access.**

**Researcher requires:** Perpetual access, Curation, Training, Marketing  
**Supplier must ensure:** Security-Access, Authorization, Authentication

---

The e-Research components as identified by the SARIS project team (Page-Shipp, et al., 2005) are inclusive of the following:

- **Data Transfer and Computation:** The infrastructure should make it possible for researchers to transfer and share, between geographically distributed researchers or groups, large data-streams or datasets, and to share models and even computing capacity.

- **Tools and applications:** Finding the available tools and applications and making them visible and available to other researchers would require efficient support and the promotion of open source and open standards.
- Primary Data Sharing: Making primary research data available to other researchers is part of the e-Science paradigm. One needs to store not only the database *per se*, but also sufficient metadata to enable a potential user to find relevant data and be satisfied with its value and provenance.

- Digital curation and preservation. This activity refers to the active management of datasets for their scientific and scholarly useful lifetimes, including the promotion of effective and widespread use. Other ePrint archives, for example, theses and research reports, form part of this category.

- Commercial and open access to eContent. The pursuit of high quality research requires ready access to the published work and data of other researchers, and the facility to publish one's own findings. Increasingly, this takes place via the World Wide Web. In the traditional commercial model, payment of a licence or other access fee is required, whereas an emerging model is of 'free' or 'open' access to online publications in formal journals or institutional or national repositories.

The basic components identified do not specifically indicate those systems or the infrastructure required to bring about the necessary change for e-Research; however, it can be seen that information, how it is managed, curated, disseminated and transferred, will be a key to e-Research systems.

The model also does not incorporate e-Learning; an essential component of skills development and most importantly does not address collaboration and collaborative knowledge development – the base reason why e-Research is changing the research paradigm. Lastly it appears that the model does not fully grasp the extent of the computing resources required to make e-Research happen.

With the size and scope of changes required to implement e-Research, it will naturally create pressure to use, whenever possible, computing
resources that do not require dedicated hardware and personnel investments. The size and scope of e-Research is vast and is composite of many disciplines and aspects. These various aspects play an important role in the successful implementation of an e-Research strategy.

Systems to support these trends are essential to bringing about the e-Research change required. The underlying concepts behind e-Research according to Hellmers (2009) are:

- Information management, retention and sharing;
- Research methods, tools and services;
- Research collaboration and dissemination.

According to Stapleton (2007), the main information management processes include gathering, analysing, communicating and storing of information. The e-Research components would need to reflect processes or infrastructure to incorporate at least these aspects of information management.

2.2.3.2 Multi-disciplinary Knowledge Management

E-Research should however be broader than just information management. Many countries are funding e-Research in order to build knowledge in their countries. However, to build a knowledge economy, calls for the integration of many disciplines, including not only information management, but also knowledge management (Mbananga, 2007).
Having information, storing it and communicating it, does not necessarily translate into knowledge generation. Managing (which would include storing) knowledge can also play an important part in e-Research to ensure that knowledge is generated, evaluated and managed correctly.

By comparing e-Research to information management requirements, and a knowledge economy model by Mbananga (2007), the following themes arise:

- The need for ICT infrastructure, tools and applications;
- The ability to transfer and store large volumes of data across networks;
- Knowledge management;
- Open access to information (or e-information);
- Stakeholder collaboration & communication;
- Information management, gathering, analysing, storage and communication.
The knowledge economy model highlights the need for knowledge management prominently; however this is not highlighted in the information services model, discussed in section 2.2.3.1. The knowledge economy model on the other hand does not specifically address the researcher’s needs for tools and services, such as curation and preservation.

Systems, tools and applications will be required to be supportive of these various needs. This raises the question as to whether e-Research can be implemented successfully when only considering those basic components of e-Research as indicated by Page-Shipp, et al., (2005).

It is suggested here that a more integrated approach is needed to support the various aspects making up the e-Research paradigm. Cyberinfrastructure, the equivalent of e-Research in the USA, appears to have looked towards a more integrated approach and is discussed below.

2.2.3.3 Cyberinfrastructure Services

Atkins (2005) in the integrated Cyberinfrastructure Services Model integrated many aspects highlighted as forming part of cyberinfrastructure.

![Figure 5: An Integrated Cyberinfrastructure Services Model (Adapted from Atkins, 2005).](image)
According to Atkins (2005), the base technologies making up cyberinfrastructure are networking, computation, storage, communication, operating systems and middleware.

The model highlights the need for certain services to allow researchers to do their work using the ICT infrastructure effectively. These services indicated by Atkins (2005) include services for observation, measurement and collaboration. Atkins (2005) further highlights the need for interfaces and visualisation services to allow users to access the necessary tools and applications.

This model does not indicate the need for open access to information, as the previous models discussed in sections 2.2.3.1 and 2.2.3.2 do. It highlights ICT infrastructure and computation services prominently, however does not show issues such as the skills development, necessary for developers of this infrastructure, support staff and the researcher using this infrastructure.

In order to have a holistic picture of e-Research, each of the aspects highlighted in the sections above, would need to be integrated into a framework, which institutions could use as a guide to the strategy development and implementation.
2.3 Summary

This chapter highlights the need to look at intangible aspects, especially in terms of bringing about change, such as a change leader. It further highlights that implementing e-Research brings about new patterns of work, and this may require skills development and transfer.

The chapter indicates that systems play a vital role in the e-Research paradigm and further points out that more research is needed to explore what exactly those systems are. It is further highlighted that structure, especially in terms of leadership, can be important to implementing e-Research.

The chapter ends with a review of the basic components of e-Research, as described by the SARIS project (Page-Shipp, et al., 2005), as well as having looked into the multi-disciplinary knowledge management model, and the integrated cyberinfrastructure model, and their applicability to e-Research. Applying these models highlighted the need for an integrated model composite of various components from the three models.

Chapter 3 provides a comprehensive literature review of the major areas addressed in the investigation. The chapter highlights those aspects that authors indicate as being components of the e-Research paradigm.
CHAPTER 3 : LITERATURE REVIEW

Whether it’s e-research in Australia, cyberinfrastructure in the United States, the grid in Europe, or e-science in the United Kingdom, a transformation is clearly occurring in research practice, a transformation that will have a profound impact on the roles of information professionals (O’Brian, 2005). Research is becoming more multidisciplinary, more collaborative, and more global (O’Brian, 2005).

The following section reports on a review of available literature and identifies various aspects/components that authors, refer to when discussing the e-Research paradigm.

3.1 e-Research Infrastructure & Systems

As a result of advances in communication technology and the internet, different networks of workstations with various capabilities have been connected to each other to form distributed computing systems (Al-Fawair, 2009), for example, the internet or World Wide Web. This emergence of distributed computing systems forms the basis of those components that make up the e-Research paradigm. Kim (2004) indicated that the e-Research paradigm is dependant on an effective ICT infrastructure/distributed systems that need to work together in an integrated way, with adequate security and perpetual and pervasive access ensured.

The European Strategy Forum on Research Infrastructures (ESFRI, 2008) indicated that the major components of e-Research infrastructures are communication networks, distributed grids, high performance computing facilities and digital repositories. These components interact with many interdependencies and as such these components are often viewed as a fully integrated system (ESFRI, 2008; Kim (2004).
Distributed grids, as referred to by the ESFRI (2008) have the potential to connect millions of computers from all over the world that are owned by many different people and organisations. By connecting all these computers, essentially one “huge and super-powerful computer” is being created (South African National Grid, 2010). As one of the fundamental concepts of e-Research is to allow researchers to collaborate, grids can be the infrastructure which allows this collaboration.

### 3.1.1 Grid Computing

In e-Research, when discussing infrastructure, many authors refer to grid computing, cyberinfrastructure or e-infrastructure (Foster and Kesselman, 2004; Ahmad, 2008; NRIC, 2006). Foster and Kesselman (1999) attempted a definition of grid computing. They wrote that a grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities. Paterson, Lindsay, Monotti & Chin (2007), however, describe grid computing as a computing model which uses the resources of many separate computers connected by a network to obtain higher levels of computational power and data processing. On the other hand, Ahmad (2008) broadens this description to say that grid computing is the uniting of pools of servers, storage systems, and networks into a single large system, to deliver the power of multiple systems resources to a single user point for a specific purpose.

Grid computing has the potential to provide on-demand access to large quantities of computing power and is also focused on large scale, often global computing, which makes it an enabling technology for e-Research (Mustafee, 2010). Grid infrastructure is beginning to underpin the operation of dynamic, virtual organizations in research, government and business (NRIC, 2006).
The basic grid technologies/components, according to the GridCoord Consortium (2006) are:

• grid middleware that enables secure, authenticated and accountable access;
• grid middleware architectures for robust, resilient and quality-of-service based services;
• grid application technologies (e.g. brokerage, work flow, indexing etc);
• grid user access technologies for grid resources (e.g. portals, portlets, etc).

This collection of technologies can be configured in different ways to serve various purposes. According to Houghton, Steele & Henty (2003), grids can be configured depending on the type of service required. This is often termed, a service orientated architecture. They further indicate that there can be the following types of grids in e-Research:

• Computing grids;
• Data grids;
• Instrument grids;
• Collaborative Working Environments; and
• Cooperative Visualization Environments.

Chhajed, Shinde, Jagtap & Raskotwar (2010) elaborate on these types of grids to include multipurpose grids, where the infrastructure of this grid should be adaptive enough provide any of the grid models.

Foster and Kesselman (2004) indicated that no matter what technologies make up the grid infrastructure or how it is configured, the main goals of the grid, should be:

• The provision of a framework for utilising unused computational power as well as unused storage capacity;
• The simplification of collaboration between different organisations by the provision of direct access to computers, software and storage;
• The provision of access to resources that cannot be accessed locally (known as “remote resources”);
• The faster execution of jobs due to their parallel execution on multiple resources (Multi-site computing/Co-allocation).

There are many grids globally. In Australia, the APAC National grid provides single sign on, data sharing and infrastructure to researchers (Maloney, 2006). The Australian Government provided AUS29 million for the second phase of the APAC grid development and in partnership with institutions such as the Australian Centre for Advanced Computing and Communications and the Australian National University have built a powerful grid architecture to further e-Research (Maloney, 2006).

In the UK on the other hand, the Parliamentary Office of Science and Technology (2007) indicated that two of the largest grids in the UK are GridPP and the NGS. GridPP received £65m in funding from the Science and Technology Facilities Council, to run a grid for particle physics. This grid is made up of 17 sites and over 5000 computers and is part of a larger international grid that will analyse data from Europe’s next particle accelerator, the Large Hadron Collider (Parliamentary Office of Science and technology, 2007). The second largest grid, the NGS is the UK’s grid for academics from any discipline; with around 500 users from over 30 institutes, this grid is funded by the JISC, which is largely responsible for university computing (Parliamentary Office of Science and technology, 2007).

The Parliamentary Office of Science and Technology (2007) indicated that there is not only a need for national grid infrastructure, but also project specific infrastructure, such as is provided by GridPP for particle physics.
South Africa has also invested in building grid infrastructure and has established the South African National grid. The South African National grid is a federation of resources, institutes and Virtual Organisations. It is a platform for collaboration and research, providing many services to scientists (South African National Grid, 2010). According to the South African National Grid (2010), the national grid form part of an e-Infrastructure, and provides support to users of scientific computing and collaboration. The distributed computing facilities are owned, managed and operated by a federation of universities, national laboratories and research groups, integrated with a middleware layer to provide seamless access to a diverse set of research communities (South African National Grid, 2010). This project is coordinated from the Meraka Institute, as part of the cyberinfrastructure programme for South Africa.

It can be seen that a distributed computing model, like grid computing is essential to facilitate e-Research. Mustafee (2010), however, argues that a new model for distributed computing is emerging, called “cloud computing”, which he posits can also provide services that are scalable through on-demand provisioning of computer resources. He further indicated that, like grid computing, cloud computing holds the promise of being a facilitator for e-Research.

Cloud Computing is hinting at a future in which we won’t compute on local computers, but on centralized facilities operated by third-party compute and storage utilities (Foster, Zhao, Raicu & Lu, 2008). There appears to be little consensus on how to define cloud computing. Foster, et al., (2008: 1) however, indicate that cloud computing can be described as:

“A large-scale distributed computing paradigm that is driven by economies of scale, in which a pool of abstracted virtualized, dynamically-scalable, managed computing power, storage, platforms,
and services are delivered on demand to external customers over the Internet”.

Cloud computing is a relatively new technology and according to Myerson (2009) to get cloud computing to work, three things are needed: thin clients, grid computing and utility computing. Myerson (2009) further indicated that with grid computing, one can provision computing resources as a utility that can be turned on or off, while cloud computing goes a step further with provisioning resources on-demand. Sundaram (2010) agrees with Myerson (2009) and elaborates further to say it is quite feasible to have a cloud within a computational grid or have a computational grid as a part of a cloud.

Whether grid or cloud computing is used as the base infrastructure, or whether a combination of the two models is used, both these technologies rely on connectivity and as such networks play an important role in building grid architecture, to enable e-Research.

3.1.2 Networks

The presence of a robust communications network is fundamental to research endeavours (NRIC, 2006). Most developed countries are investing in upgrading their research and education networks (NRIC, 2006). Countries such as the UK and Australia have made vast investments to improve existing networks to allow e-Research.

In Australia and the UK, various networks make up the network infrastructure to provide e-Research capabilities. Both Australia and the UK have Research and Education Networks. In Australia, the network is commonly referred to as AREN, while in the UK it is called Janet (Australian Government, 2006; JISC, 2011).

The Aren network links major metropolitan cities, most universities’ main campuses, and a link across the Pacific interconnecting with
other high-speed academic and research networks in the US and through them with Europe and Japan (Australian Government, 2006). The Australian Government has invested more than $80 million in this network’s development and improvement (Australian Government, 2006).

According to the Platforms for Collaboration (2010), the existing network infrastructure for research is insufficient, specifically in connecting certain regional areas. They further indicate that for e-Research to be truly effective, more investment will be required to improve connectivity internationally. As shown in Figure 6 below, many regional areas are not connected to the Aren network. No indication could be found that the network infrastructure in Australia has changed vastly since the publication by Sargent (2002).

![Figure 6: Conceptual Aren Backbone (Sargent, 2002)]
In the UK, the main network interconnecting the main cities is commonly referred to as Janet. Janet (UK) is government funded, with the primary aim of providing and developing a network infrastructure that meets the needs of the education and research communities. It connects the UK’s education and research organisations to each other, as well as to the rest of the world through links to the global Internet. In addition, Janet includes a separate network that is available to the community for experimental activities in network development.

The Janet network connects UK universities, FE Colleges, Research Councils, Specialist Colleges and Adult and Community Learning providers (Janet, 2010). It also provides connections between the Regional Broadband Consortia to facilitate the Department for Education and Skills initiative for a national schools’ network (Janet, 2010). Janet also brings together regional networks such as Clydenet, the Clyde area network and Niran, the Northern Ireland Regional Area Network.

*Figure 7: SuperJanet4 Backbone (Hey and Trefethen, 2003)*
According to the Australian Government (2006), the existing network infrastructure for research is insufficient specifically in connecting certain regional areas. The focus of building networks was on connecting major cities. However, in the UK, it was found that many regional networks were established early on in the development of e-Science (Janet, 2010). However, institutions were not all fully integrated, due to the lack of national and international networks. From the UK and Australian examples it can be seen that both regional networks and inter-institutional networks, for example linking all universities in a region together, which then connect to a national network, may be necessary to enable e-Research. As shown in Figure 8 below.

Figure 8: UK Regional and National Network Architecture (Sargent, 2002)
Networks to facilitate e-Research should have high bandwidth, low latency and low jitter (Australian Government, 2006). They must also be reliable, support mobility, be connected to local communities and interoperate with international peer networks (Australian Government, 2006).

One of the technical challenges to be overcome relating to networks, is the ability to enable research teams to construct ‘virtual research environments’ on demand, and adapted to particular projects (Australian Government, 2006: 32). In this regard, Virtual Private Networks will become increasingly important as collaborative research methods increase (Australian Government, 2006). This requires that components of the e-Research infrastructure and researchers can interoperate with one another, which is usually permitted through network infrastructures.

In South Africa, the National Research and Education Network (SANReN) was established to link all major metropolitan cities, like Johannesburg, Cape Town and Durban together (SANReN, 2011). It was found that, like the UK and Australia, all metropolitan cities have been linked to a high speed backbone network. SANReN (2011) has within these cities further established a smaller network, linking research and education institutions together (inter-institutional networks).
SANReN forms the backbone for connectivity into the rest of Africa and across the world, linking South Africa to research and education networks globally (SANReN, 2011). This project is being funded by the Department of Science and Technology (DST), through the CSIR Meraka Institute (South African National Grid, 2009).

According to many authors, networks alone are not sufficient for building grid capabilities in countries (Burton, 2007; Sargent, 2006; ESFRI, 2008). Burton (2007) indicated that some of the elements of e-Research infrastructure are well-known and well-established. One of these elements, according to Burton (2007), is high performance computing facilities.
3.1.3 **High Performance Computing**

The demand for high performance computing systems with increased capacity and capabilities has been traditionally driven by the need to model and simulate complex natural systems and processes in, for example, chemistry, physics, biology, geology and the environment (NRIC, 2006). There are an increasing number of users/researchers for whom access to large-scale data is an essential requirement of their research. They are often concerned with data processing techniques such as searching, filtering, comparing, mining and pattern discovery (NRIC, 2006).

To meet these requirements, researchers now need access to powerful high performance computing capacity, mass data storage systems, interactive visualization systems and high capacity communication services (NRIC, 2006). They require services such as grid computing and access to databases that make increasingly more extensive use of high performance computing facilities in a collaborative environment (NRIC, 2006). Consequently, the demand for high performance computing can be expected to increase over time.

When investing in high performance computing systems, the (ESFRI, 2008) highlights that there is a fundamental difference between capacity and capability in terms of computing resources. While grid infrastructure meets the needs for capacity, a different aspect is capability computing. In this case, not only the power and/or the number of processors are relevant as in grid, but also the system architecture and especially the size of the core memory and the throughput bandwidth between the computing engine and the memory are of vital importance for performance.

In Australia, a number of high performance computer centres have been established, all connected to the research network (Australian Research Council, 2004). These facilities include: large computer
clusters in Canberra, managed by Australian Partnership for Advanced Computing (APAC); in Melbourne, managed by Victorian Partnership for Advanced Computing (VPAC); in Sydney, managed by Australian Centre for Advanced Computing and Communications (ac3); in Brisbane, managed by Queensland Parallel and Supercomputer Facilities (QPSF); and a number of smaller computer server installations such as in universities (Australian Research Council, 2004). Many of these computing facilities are networked and therefore can provide distributed computing capability. According to the Australian Research Council (2004), it is important that as many such high performance computing facilities as possible be available to service researchers’ increasing advanced computing needs.

While in Australia focus has been on establishing high performance computing centres across different states, within the UK focus has been on building Hector. Hector is the UK’s largest, fastest and most powerful supercomputer (Hector, 2007). It is capable of over 470 “million million” calculations a second, which is over 68,000 calculations a second for every man, woman and child on Earth (Hector, 2007). It also has one Petabyte of disk space for storing data. Funded by UK Research Councils, and operated by various institutions appointed by the UK Research Council (Hector, 2007).

Hector allows researchers in the UK to tackle today's biggest scientific and engineering challenges. Activities that rely on modern high-end computing tools and techniques include: forecasting climate change, designing new life-saving drugs, constructing safer aircrafts, predicting natural disasters, and understanding how complex biological systems work and develop (Hector, 2007).

Like the UK, South Africa has focused on developing one site for high performance computing, the Centre for High Performance Computing (CHPC, 2010) in Cape Town. In 2009, the centre launched “Sun Constellation System” which is proudly listed as the fastest
supercomputer in Africa (CHPC, 2010). Through the use of this system the CHPC aims to “enhance significant research, address grand challenges, and grow computational research” (CHPC, 2010). This project is being funded by the Department of Science and Technology (DST), through the CSIR Meraka Institute (CHPC, 2010).

From the literature reviewed in sections 3.1.1 to 3.1.3, the various authors have highlighted that possibly some of the basic components making up the e-Research framework, would include networks and high performance computing, as shown in Figure 10 below.

![Figure 10: Basic ICT Infrastructure for e-Research](image)

Advanced ICT or grid infrastructure comprises more than computer hardware, associated software and telecommunications networks that are used in standard office environments. In particular, it includes a new level of software tools and services, collectively termed ‘middleware’, that link the ICT resources that users need (Australian Government, 2006).

### 3.1.4 Middleware

A major development of grid computing has been the development of middleware, the communications layer that allows applications to interact across hardware and network environments (Paterson, et al., 2007). The grid middleware is software that provides standard community tools and services for knowledge management, knowledge sharing, collaboration and interoperability between applications, computing resources, institutions and individuals across the grid (DSTC, 2004; Byeon, Kwon, Dunning, Cho & Savoy-Navarro, 2009).
The grid infrastructure is built with this set of grid middleware, which work together to provide transparent resource sharing environment for upper application (Li, Yang, Jiang & Shi, 2006). Middleware allows researchers and applications to treat the data repositories, computing, and other disparate resources as if they were one large virtual facility. There have been many distinguished middleware platforms, such as Globus Toolkits (Foster & Kesselman, 1999) and Legion (Grimshaw, Wulf, & James, 1994) amongst others. The success of these platforms makes it possible to build new grid infrastructure for wider resource sharing (Li, et. al., 2006).

A number of large ‘production-orientated’ projects around the world are developing and deploying middleware (DSTC, 2004). These are cooperating under the auspices of the Global Grid Forum to harmonise their developments and enable grid interoperability (DSTC, 2004). According to the DSTC (2004), the goal is to create an operational grid infrastructure, capable of supporting cooperation by multi-national research teams. A significant achievement is the Globus Toolkit, a key component of most production grids (DSTC, 2004; Byeon, et al., 2009). Byeon, et al., (2009) indicated that the Globus Toolkit is used to construct basic grid middleware services such as program execution and file transfer.

Many current e-Research projects, in Australia, the UK and internationally, are being undertaken by teams working largely independently. However, many of these projects could have common middleware needs. According to Sargent (2006), they all require seamless access to e-Research resources in order to access the services offered by those resources.

In Australia, the Middleware Action Plan and Strategy (MAPS) Project was established to guide researchers, universities and other research institutions in the adoption of a common middleware strategy,
standards and tools (Sargent, 2006). In the UK, the same was done through the establishment of the Open Middleware Infrastructure Institute.

According to the DSTC (2004), the UK e-Science initiative is composite of three main components,
1. An Open Middleware Infrastructure Institute (OMII); and,
3. A number of Regional Centres
The OMII and GSC work in close partnership to facilitate the acquisition, coordination, development and integration of the middleware required by e-Research communities and to support pilot projects and communities in the Regional Centres which use the middleware and Grid environments, within the UK (DSTC, 2004). This could also possibly work in the South African context.

The DSTC (2004) elaborates on the various middleware developments taking place and indicate that middleware can be classified into three main categories, being:

- Grid Services/Resource Management middleware, which includes the Open grid Services Infrastructure (OGSI), provides essential access, communication, accounting, security, trust, and co-ordination services between the computational and data resources of the grid and the higher-level services that use them.

- Knowledge Management (KM) middleware provides tools and services that enable the indexing, archival, discovery, analysis, integration, management and preservation of large heterogeneous distributed data repositories and digital archives.

- Collaborative middleware provides tools and services to support formal and informal, real-time and offline collaborative activities between remotely located researchers, research communities, and resources (dynamic, scalable virtual organizations).
While the DSTC (2004) incorporates in its description of middleware, certain tools for collaboration, collaboration is central to the need for e-Research, and as such, is addressed as a set of tools on its own as well as being a part of the middleware.

E-Research infrastructure has been conceptualised largely due to increasing volumes of data or the “data deluge”, and as such, Sargent (2002) and Wilbanks (2009) indicated that managing this data effectively is essential to supporting the full cycle of research endeavours, from research concept formulation and scoping, to the research activity itself, to the dissemination of the research results.

3.2 Managing Data, Information and Knowledge

Large-scale e-Research environments have to deal with numerous challenging and crucial data management issues (Fiore and Aloisio, 2011; Wilbanks, 2009). Application domains where data management is becoming more and more challenging span among others, High Energy Physics, Earth and Environmental Sciences, Bioinformatics, Astronomy and Astrophysics (Fiore and Aloisio, 2011).

According to many authors (Androulakis, et al., 2009; Wilbanks, 2009 & Parastatidis, 2009), the need for the data management infrastructure is being felt ever more acutely in the e-Research community due to:

• The quantity of scientific data increasing exponentially, challenging researchers to keep track of it all;

• This large quantity of electronic data creating new challenges for collaboration;

• Much greater expectations placed on online publishing of data and verifiability of experiments;

• Concerns about security and privacy in many disciplines of e-research; and
• A significant push to streamline the workflows of e-research by providing centralised, persistent, and reliable storage.

Through the use of technology and automation, institutions are trying to keep up with the challenges of the “data deluge” (Wilbanks, 2009; Parastatidis, 2009). Parastatidis (2009) indicated that computers and software have become efficient at storing, managing, indexing and computing data. Sargent (2002) further indicated that data management needs of researchers were identified as being comprised of:

• data collection and generation;
• data storage and the physical management of stored data;
• evolution of standards and protocols to facilitate the storage, use and interpretation of data;
• access to data;
• Long-term archival and preservation of data and policy for retaining and discarding of data.

ICT can assist in meeting the needs of researchers, especially in terms of storage and transfer. Parastatidis (2009), however, indicated that we are a long way from having computer systems that can automatically discover, acquire, organise, analyse, correlate and interpret large volumes of data, especially on a global scale.

To address these data management needs, the Australian Government, via the Department of Education, Science & Technology, established the Archer project (Androulakis, et al., 2009). The Archer project has produced a suite of tools, developed jointly with universities in Australia, to assist researchers in collecting, managing, storing, collaborating on, and publishing scientific data (Androulakis, et al., 2009).
Similarly, in the UK, the e-Science Data Management Group was established to provide an improved framework for data management, curation and preservation (Science & Technology Facilities Council, 2010). Research is being done by both the e-Science Data Management Group in the UK and by the Archer Project, into metadata, data access technologies, security of data, security architectures, and various other aspects of data management, such as curation (Androulakis, *et al*., 2009; Science & Technology Facilities Council, 2010).

Parastatidis (2009) indicated that the primary focus of the current technologies addresses only the first part of the data-information-knowledge-wisdom spectrum. While technologies are focusing on data and information management, Parastatidis (2009) indicated that knowledge management in relation to e-Research is becoming more important, in order to act as a foundation for a generation of knowledge-driven services and applications. He further highlights the need for knowledge orientated research infrastructures, especially in light of the fact that countries are using these infrastructures in the hope of building knowledge-based economies.

Mohamed (2007) indicated that e-Research infrastructure builds on grid computing, in that it builds a knowledge environment through the addition of higher mechanisms to facilitate knowledge assimilation, transformation, mobilization, and sharing. He further indicated that the focus of many countries was on ground-breaking collaborative discoveries and grid computing, while improving network speeds, storage capacity and throughput of systems. He indicated, however, that aspects such as knowledge management have not been in the forefront. Mohamed (2007) does, however, indicate that the components of e-Research go a long way towards building infrastructure for mobilising knowledge in the global knowledge-based economy.
In terms of Data, Information and Knowledge Management, authors highlight three main aspects; that of storage of information – digital repositories, having access to information to enable research, and digital curation and preservation of the information (Androulakis, *et al*., 2009; Page-Shipp, *et al*., 2005).

With the increased production of data through modern research activity, and the use of new research infrastructure, and with the outputs from simulations and the various instruments and sensors among the various research communities, infrastructure providing very large storage capacity is required to store and make accessible key research data (Androulakis, *et al*., 2009). Digital repositories play a significant role in the storage of research data and information (Androulakis, *et al*., 2009).

### 3.2.1 Digital Repositories

In simplest terms, a digital repository is where digital content is stored and can be searched and retrieved for later use. A repository supports mechanisms to import, export, identify, store and retrieve digital assets (JISC, 2011). Putting digital content into a repository enables staff and institutions to then manage and preserve it, and therefore derive maximum value from it (JISC, 2011). According to Francis (2006a), there are various types of digital repositories, for example, research databases and online libraries.

Data in their various forms, from raw data to scientific publications, will need to be stored, maintained and made available and openly accessible to all scientific communities (ESFRI, 2008; Page-Shipp, *et al*., 2005). Presently, there are a myriad of efforts to store and manage research data, largely based around institutions and, within institutions, around departments and individuals (ESFRI, 2008). The quantity of research data is growing rapidly. Investment is needed to ensure that this data is managed in a coherent manner, so that it can be readily accessed when and as it is needed (ESFRI, 2008).
ESFRI (2008) indicate that in many cases the data is managed through middleware tools and applications.

Repositories provide much greater functionality and support for collaboration, as well as exposing research activity in ways that have not been possible before, further increasing the return on investment in publicly funded research (JISC, 2011). Digital repositories may include research outputs and journal articles, theses, e-learning objects and teaching materials or research data (JISC, 2011).

In South Africa, Van Deventer and Pienaar (2008: 3) indicated that there were many institutional repositories, giving access to “at least nine open access repositories collections at several academic institutions”. These repositories were initially focused on open access to electronic theses and dissertations, however the focus is changing and “more institutions are now investigating the possibility of making their data collections accessible” (Van Deventer and Pienaar, 2008: 4).

The JISC (2011) indicated that repositories are important for universities, colleges and research institutes in helping to manage and capture intellectual assets as a part of their information strategy and further to contribute towards building knowledge economies. In addition, they indicate that a digital repository can hold a wide range of materials for a variety of purposes and users. It can support research, learning, and administrative processes.

In South Africa, the “Very Large Database (VLDB)” is in the development phase (CHPC, 2010). It is aimed at complementing the Centre for high performance computing through the effective curation of large databases, specifically in areas of environmental and climate change modelling and astronomy (CHPC, 2010). This project is being funded by the Department of Science and Technology (DST), through the CSIR Meraka Institute (CHPC, 2010).
In all cases, the provision and maintenance of institutional repositories, the software, hardware and services required to accept, store, make available online and manage a wide range of digital content, including the research output – are fundamental to e-Research (JISC, 2011).

While researchers can have access to the necessary infrastructure and appropriate tools and application, without access to information, comprehensive research cannot be undertaken.

3.2.2 Access to Information
Scientific and scholarly publishing is now evolving along two distinct paths – one in which large commercial publishers are increasing their dominance in such areas as ‘branded’ journal titles and access to scientific publications, and the other in which there are a variety of open access initiatives (Houghton, et al., 2003).

According to Houghton, et al., (2003), open access digital repositories, operating in parallel with existing commercial publishing mechanisms, may provide a major opportunity to develop a sustainable information infrastructure for both traditional and emerging modes of knowledge production. They further indicate that together, they provide the foundation for more effective and efficient access to, and dissemination of, scientific and scholarly information.

Recognition of the significance of information and knowledge has led to an appreciation of the economic and social importance of access to Information (Paterson, et al., 2007). The move towards “open access” of scholarly journals, digitisation of library materials, and universities opening their access to research, in the form of thesis and dissertation material, gives researchers in all fields access to more information than was previously available (Mutula, 2009). The open access movement is described as being composed of two components, being
the open access to journals and the open access to archives and repositories (Taylor and Weaver, 2005; Mutula, 2009).

Currently extensive holdings of research data are stored within personal archives, either on researcher desktops or on departmental/institutional servers (Australian Government, 2006). In these locations, it is largely inaccessible, and inhibits collaborative research activity.

Data generators, such as scientific facilities, and data holders, such as museums and archives, are rapidly increasing their generation and holdings of digital data (Australian Government, 2006). Researchers are increasingly relying on digital data where appropriate, due to its advantages for research analysis (Australian Government, 2006). The e-Research consultation process around Australia highlighted the critical need for researchers to have access to data, whether it is scientific data generated by scientific facilities, or humanities and social sciences data from historical records, and whether it is held in Australia or overseas (Sargent, 2007). Research outputs are highly dependent on the input data, implying that researchers need to be confident of the accessibility, persistence and integrity of the data they need to access (Sargent, 2007).

The linking of distributed e-Research resources, such as scientific facilities, digital data repositories and computing capability with researchers, brings with it responsibilities that data accessed in such a way is not only easily accessible, but also that its persistence and integrity are assured so as to meet researchers’ needs.

Houghton, et al., (2003) & Patkar and Chandra (2006) indicated that researchers voiced concerns about open access to information. They further indicated that among the researchers’ concerns was the concern that the quality of information online was of such a nature that it could not be used in research. Houghton, et al., (2003:118)
indicated that one researcher said that “the quality of information online is rubbish”. Patkar and Chandra (2006) indicated that besides quality the two most critical integrity issues with web-based information are: provenance - where does an object come from and how has it changed over time, and persistence - how long will an object last and how to make it last longer.

This highlights the need for researchers to have access to good quality information, but also the importance of curation and preservation of the information.

Open access to information is, however, about more than just where a country is in terms of repositories or open access journals (Fonseca, 2009). It helps researchers and scientists work globally, in a more cooperative environment, and in more balanced terms (Fonseca, 2009). Fonseca (2009) indicated that it was proven by research that open access provides faster and greater visibility to research, especially in developing countries. He further indicated that these elements in turn help the communication and scientific communication processes, and help build knowledge from previous experience and research.

To enhance researcher effectiveness and facilitate easier access to research results and outcomes, it is also essential that electronic storage of research is consistent with internationally agreed technical standards (Fonseca, 2009).

Sargent (2006) indicated the e-Research Coordinating Committee identified easier access to data by researchers as a critical issue in Australia. He further indicated that access to information has become an issue due to a variety of factors including the various standards under which the data is presented, the lack of description under which the data is collected, the lack of proper access permission provided to the data collected, and the data being stored offline.
Francis (2006b) stated that the NCRIS identified two different types of data being, private and public data. He indicated that private data is usually held by a researcher through self generation or by acquisition from surveys, an instrument or device. Public data is data which has been placed in a public repository for general public access.

Borgman (2006) agrees with Francis in saying that primary data is anything raw and unprocessed, such as sensor data and field notes. Borgman (2006) elaborates to include secondary data, which she defines as data that has been analysed, processed or interpreted in some way. Various authors have indicated that access to information is inclusive of access to primary data, access to open publishing and access to commercial publishing.

Having access to information means that information must be available to be accessed; this highlights the need for digital curation and preservation (JISC, 2011).

3.2.3 **Digital Curation & Preservation**

Data curation came to light as being essential to the e-Research agenda in early 2000 (Lord and Macdonald, 2003). The recognized importance of data curation in e-Research relates to an awareness of the exponentially increasing volumes of primary data in digital form generated by automated collection of data through “next generation” experiments, simulations, sensors and satellites (Hey and Trefethen, 2003), that need to be archived and preserved, so that they are available and appropriate for contemporary discovery and future reuse (Lord and Macdonald, 2003).

Hockx-Yu (2006) refers to work done by Jones and Beagrie (2002), who indicate that digital preservation refers to the series of managed activities necessary to ensure continued access to digital materials for as long as necessary. The sheer volume of data in digital repositories is a challenge for the preservation of the data (Hockx-Yu, 2006). In the
UK, the OSI e-infrastructure Working Group (2004) indicated that digital data and information do not survive without active intervention. They further indicate that mechanisms to ensure long-term preservation are vital to the research lifecycle.

In the research context, major reasons to keep primary research data include (Lord and Macdonald, 2003):

- Re-use of data for new research, including collection-based research to generate new science;
- Retention of unique observational data which is impossible to re-create;
- More data is available for research projects;
- Compliance with legal requirements;
- Ability to validate research results;
- Use of data in teaching;
- For the public good.

In the UK, the JISC (2011) established the Digital Curation Centre (DCC) which focuses on building capacity, capabilities and skills for data curation across the UK. In Australia, curation of data is highlighted, not as a separate process, but rather as forming part of the data storage processes (NRIC, 2006).

From the literature reviewed in section 3.1.4, authors have highlighted that middleware is necessary as a communications layer that allows applications to interact across hardware and network. From the literature reviewed in section 3.2, authors have highlighted that Data and Information Management are important in e-Research, as well as having access to information and digital repositories, while ensuring that data is curated. These components indicated can thus possibly also form part of the components making up the framework for e-Research as shown in Figure 11 below.
Curation and preservation are highlighted as being important components of e-Research; however, authors also highlight that there are other tools and applications which can play an important role in implementing the complete e-Research paradigm.

### 3.3 Tools and Applications

Different users, disciplines and applications have traditionally made their own arrangements and written suitable software and systems to achieve various tasks (Sargent, 2006). As the numbers and type of distributed electronic resources has increased, and as collaborative ways of working have proliferated, it has become evident worldwide that this individual approach could be wasteful, as it leads to an expensive duplication of resources and an increased risk of developing systems that cannot communicate easily with each other (Sargent, 2006).

Sargent (2006) indicated that especially within the research community, and now increasingly among the general academic community, a substantial amount of effort is being expended in various international working groups to identify common standards and practices that may eliminate much of this duplication, and to develop
systems that can be applied across a wide spectrum of applications and resources.

As Sargent (2006) states, one of the main focus areas for the development of tools and applications, has been those tools that allow researchers to communicate and collaborate.

3.3.1 Communication and Collaboration

Effective collaboration between researchers, including shared access to repositories and related research infrastructure, is intrinsic to the e-Research vision (Sargent, 2006, 2002, 2007; O’Brian, 2005, Francis, 2006b; NRIC, 2006 & Cummings and Kiesler, 2005). Collaboration may involve remote control of equipment, sharing of data, and access to content in repositories and increasingly, communication and conferencing (Sargent, 2006), including audio conferencing, video conferencing and instant messaging.

Collaboration, especially international collaboration, is widely seen as beneficial. Walsh and Maloney (2002) found that computer network use is associated with more geographically dispersed collaborations as well as more productive collaborations. Katz and Hicks (1997) found that the citation of collaborative papers was higher than that of sole authored papers, and that the impact was higher for collaborative papers where those collaborations were international. These findings suggest that collaboration may produce increases in both productivity and quality (Katz and Hicks, 1997).

e-Research collaboration tools have created new opportunities for research and communication among individuals, particularly for those that are geographically dispersed (Cummings and Kiesler, 2005). Skype, Voice over Internet Protocol (VoIP), e-mail and instant messaging are digital tools facilitating communication and collaboration and allow people across many geographical locations to
coordinate tasks and research output (Anandaraian and Anandaraian, 2010).

The Australian Research Collaboration Service (ARCS) provides national e-Research services via grid-enabled infrastructure (Platforms for Collaboration, 2008). ARCS provides a vital platform for research collaboration including seamless access to research facilities and services, such as high performance computing resources, via interoperable middleware, as well as providing for the analysis and movement of large data sets (Platforms for Collaboration, 2008). The Platforms for Collaboration (2008) committee funds ARCS which provides national services through a cohesive collaboration of regional partners (Platforms for Collaboration, 2008).

Another important aspect of e-Research which is highlighted by many authors is that of visualisation tools and applications.

### 3.3.2 Visualisation

Visualisation can be divided into two broad types or categories, Scientific Visualization and Information Visualization (Charters, 2010; Ware, 2004). According to Charters (2010), scientific visualisation is concerned with the conversion of numeric data, usually from scientific experiments and simulations to a graphical representation. He, however, indicated that information visualisation, on the other hand, is concerned with the conversion of other forms of data, structured and unstructured text, images and video, to an appropriate graphical representation. The representations generated by the visualisation process can be abstract in nature, or can be based on a physical representation or a real world metaphor (Charters, 2010).

Butler, *et al.*, (1993) describe three categories of user oriented visualisation task:
• Descriptive Visualisation, occurs when the phenomena represented in the data are known, but when the user needs to present a clear visual verification of the phenomena;
• Analytical Visualisation, or directed search, is the process we follow when we know what we are looking for in the data; visualisation helps determine whether it is there;
• Exploratory visualisation, or undirected search, is necessary when we do not know what we are looking for; visualisation may help us understand the nature of the data by demonstrating patterns in that data.

Visualisation has many advantages for a research investigation; five of these are (Ware, 2004):
• Ability to comprehend huge amounts of data;
• The perception of emergent properties;
• Easy indication of errors in the data;
• Understanding large-scale and small-scale features;
• Facilitation of hypothesis formation.

In line with Ware’s (2004) findings of the advantages of visualisation, and expanding on these advantages, Hansen, Johnson, Pascucci & Silva (2009: 153), indicated that visual data analysis, facilitated by interactive user interfaces, “enables the detection and validation of expected results, while also enabling unexpected discoveries in science”. These advantages show visualisation has an important role to play in e-Research; however, the development of visualisation technologies has not kept pace with the changing needs of visualisation users (Ware, 2004). Hansen, et al., (2009: 153) agrees with Ware (2004) and further indicated that visual data analysis is far from being a “solved problem”. 
Visualisation highlights that for different research projects, different visualisation tools may be required. This highlights the possible need for project-specific tools and applications.

### 3.3.3 Project Specific Tools and Applications

Dozier and Gail (2009), state that for earth and environmental science specific tools and applications are required. Examples highlighted by the authors include tools for measuring rainfall and rising sea levels. They further indicated that models such as information models do not consider interactions between various variables stored in the information, and hence indicated the need for “environmental applications”. Delaney and Barga (2009) also indicate that in the field of oceanography, specific tools and applications would be needed for researchers to develop predictive models for giant storms or eruptive volcanoes.

The same can be said for the SKA Africa project. This project, the Square Kilometre Array (SKA), will focus on exploring the universe with the world’s largest telescope array (SKA Africa, 2011). The SKA will be a radio telescope, which instead of seeing light waves; will generate pictures from radio waves (SKA Africa, 2011). To capture, store and compute the information from this telescope array, vastly different tools and applications will be required than those used for studying oceanography.

It can be seen that project specific tools and applications, such as highlighted by Dozier and Gail (2009) and Delaney and Barga (2009) may be necessary in an e-Research framework.

A growing concern in e-Research communities is that of security, especially in terms of access, authentication and authorisation.
3.3.4 **Access, Authentication and Authorisation**

Supporting e-Research often places numerous demands on capabilities associated with e-Research Infrastructures (ISSR, 2010). One key challenge is with regard to security. Security technologies must support a variety of user and provider requirements such as authentication, authorization, trust, privacy, policy management, and information assurance, all in a user and provider friendly framework (ISSR, 2010). Many e-Research projects have developed distinctive security analyses, requirements, and technical solutions for a variety of applications (ISSR, 2010).

Like most complex software systems, grid middleware solutions exhibit a number of security problems (Smith, *et al.*, 2009). Unlike traditional cluster systems and the small academic grid initiatives where local administrators usually know their users’ software and usage habits, the larger mixed academic and business grids expose cluster administrators to a large number of unknown users with a great variety of usage patterns (Smith, *et al.*, 2009).

Grids are now becoming an attractive target for attackers, since the grid offers standardized access to a large number of machines storing potentially valuable data, which could be misused in various ways (Smith, *et al.*, 2009).

The JISC established the UK Access Management Federation for Education and Research, to provide solutions to some security issues (JISC, 2011). The federation is largely responsible for providing single sign-on solutions for accessing online resources and services (JISC, 2011). While in Australia, the Australian Access Federation (AAF) was established to bring together cutting-edge technologies as a framework for trusted electronic communications and collaborations between universities and research institutes (Australian Access Federation, 2011).
According to Sargent (2006), the security of e-Research infrastructure has to be addressed at all levels; it cannot effectively be addressed unless it is with respect to all elements in a communication chain. Research institutions in Australia have particular security problems born of their distinctively large bandwidth and their connectivity to global high performance networks (Sargent, 2006).

e-Research is opening up new fields, previously deemed to not be possible; many universities implementing e-Research are finding uses for its capabilities beyond just research. e-Research is broadening in its context and emerging uses are being found for these technologies.

3.4 Trends in e-Learning and Digital Scholarship

Universities across the world are discovering that implementing an e-Research strategy also allows for e-Learning or digital scholarship. As e-Research is being implemented in institutions world-wide, and as grid computing is being recognized and used, a new form of grid computing called e-Learning grid is emerging as a nationwide e-Learning infrastructure, which can provide innovative learning experience for learners (Li, et al., 2006). According to Li, et al., (2006) grids can resolve many of the existing drawbacks in e-Learning environments, such as scalability, interoperability and availability.

e-Learning or digital scholarship is described by Mutula (2009) as being a ‘networked’, scholarly or academic environment with pervasive integration of digital technologies in everyday learning. He elaborates to say it is inclusive of teaching, using electronic means to evaluate and assess academic work, e-journals, e-books and digital libraries. Many of the tools and infrastructure required by universities to make provision for digital scholarship are also required for e-Research. As Mutula (2009) indicated, digital scholarship and e-Research are closely related and often interlinked.
By using e-Research infrastructures, learning content can be made available even in remote locations (Mutula, 2009).

The University of Reading in the UK found that the use of e-Research technologies helped to advance their e-Learning capabilities (University of Reading, 2010). They found that more extensive use of e-Research resources enhanced undergraduate teaching and learning (University of Reading, 2010). They have further incorporated many schools in the area, through their Blackboard School Portals project, which provides resources to teachers, such as building lesson plans, which can be easily updated each year.

Encouraging and facilitating collaboration and coordination with e-Learning Strategy, includes sharing of resources and the use of digital research outputs for teaching. E-Learning at Reading (and nationally) currently focuses heavily on the use of Blackboard; there is scope for more extensive use of e-Research resources for undergraduate teaching and learning.

From the literature reviewed in section 3.3 and 3.4, authors have highlighted that various tools and applications may be necessary to support e-Research activities. These components indicated can thus possibly form part of the components making up the framework for e-Research as shown in Figure 12 below.

![Figure 12: Tools and Application Components for e-Research](image-url)
Many authors indicate that funding structures play an important role in the implementation of e-Research strategies. The next section looks at what authors say about structure in relation to e-Research.

### 3.5 Funding Structures & Oversight of e-Research Activities

In this literature review, structure will only be looked at in so far as authors highlight specific areas relating to structure, which have an impact of the implementation of e-Research activities. In this literature review, structure looks to what authors indicate about funding structures, as well as co-ordination and oversight of e-Research activities.

As e-Research involves often large scale collaboration infrastructure developments, funding these activities is usually larger than the scope of a single institution. The following section investigates the total investments made by countries into e-Research activities, and also looks to whether these activities are funded at a national or institutional level. This is undertaken to better understand how a framework of e-Research components may be funded in the most cost-effective way possible, learning from the experiences of developed countries.

#### 3.5.1 Overview of Global e-Research Funding

According to Alberta’s Cyber Infrastructure taskforce (2006), nations around the world have established e-Research as a priority, recognising it as essential to remaining competitive in the global research arena. They further indicate that these activities are largely supported or funded by governments, who are making significant monetary investments into e-Research initiatives, funding both initial capital and ongoing operating expenses. The figure below gives an indication of the overall investments made by various nations into e-Research (Alberta’s Cyber Infrastructure taskforce, 2006).
It can be seen that investment into e-Research activities is considerable. In order to fund e-Research cost effectively, there may be lessons to be learnt from the way in which Australia and the UK undertook funding their e-Research programs. The funding structures of both Australia and the UK are investigated in sections 3.5.2 and 3.5.3.
3.5.2 **Australian Funding Structure of e-Research Activities**

Appelbe (2010) indicated that infrastructure in Australia was funded initially very differently from the way it was done in the UK. He further indicated that early in 1995, universities and institutions supported and funded their own high performance computing and e-Research infrastructures, which resulted in very little collaboration. In 2000, APAC (Australian Partnership for Advanced Computing) was established (Appelbe, 2010). Commonwealth funding was provided to states to set up and support e-Research centres, specifically high performance computing centres. Appelbe (2010) further indicated that since 2005, a strong emphasis has been placed on collaboration rather than on competition for research infrastructure. Funding focused on setting up, for example, one high performance computing centre in each state, whereby universities and research institutions, by and large, outsource HPC activities to these state organisations (Appelbe, 2010).

Appelbe (2010) indicated that lessons learned in Australia with regards to funding e-Research activities show that partnerships are needed for funding e-Research. He indicated that these partnerships should be comprised of university/research institutions, state, regional and federal funding.

In Australia, the NRIC (National Research Infrastructure Council) was established to develop and fund national scientific infrastructure projects (NRIC, 2006). The Australian Government provided $542 million over the period 2005-2011 to provide researchers with research facilities, supporting infrastructure and networks necessary for world-class research (NRIC, 2006).

Furthermore, according to the Alberta’s Cyber Infrastructure taskforce (2006), an investment of £313 million plus £ 25 million (industry contributions) was put into the UK e-Science Initiative. These activities are indicated as being funded by government (Office of
Science and Technology; Department of Trade and Industry) and industry. While in Australia, they indicate the following investments were made:

- $246.0 million – Systemic Infrastructure Initiative (SII) between 2001-2006;
- $145.7 million – Major National Research Facilities Programme (MNRFP) between 2001-2006;
- $542.0 million – National Collaborative Research Infrastructure Strategy (NCRIS) between 2004 and 2011.

According to the NRIC (2006), major infrastructure should be developed on a collaborative, national, non-exclusive basis. Infrastructure funded through the NCRIS will serve the research and innovation system broadly. Funding and eligibility rules of the NRIC encourage collaboration and co-investment in e-Research (NRIC, 2006). The NRIC (2006) further indicated that it is not the function of NCRIS (National Collaborative Research Infrastructure Strategy) to support institutional level (or even small-scale collaborative) infrastructure.

3.5.3 **UK Funding Structure of e-Research Activities**

In the UK, the JISC – Joint Information Systems Committee was established in 1993. The committee has the responsibility to deal with networking and specialist information services, provide these services while exercising vision and leadership in bringing about developments for the benefit of the Higher Education councils for Scotland, Wales and England. The JISC further supports a range of national infrastructure services to address the needs of the Higher Education and Further Education communities in the UK. The JISC is largely responsible for the implementation of e-Research, or e-Science as it is known in the UK.
The JISC is funded by institutions such as the higher education funding council for England and the education funding council for Wales. These are largely government-funded councils.

According to Hey and Trefethen (2002), under the UK Government’s Spending Review in 2000, the UK’s Office of Science and Technology (OST) was allocated £98M to establish a 3-year e-Science R&D Programme. They further indicate that this e-Science initiative spans all the research councils— the Biotechnology and Biological Sciences Research Council (BBSRC), the Council for the Central Laboratory of the Research Councils (CCLRC), the Engineering and Physical Sciences Research Council (EPSRC), the Economic Social Research Council (ESRC), the Medical Research Council (MRC), the Natural Environment Research Council (NERC) and the Particle Physics and Astronomy Research Council (PPARC). A specific allocation was made to each Research Council (shown below). As is common in Department of Trade and Industry (DTI) programmes, the DTI contribution of £20M requires a matching contribution from industry (Hey and Trefethen, 2002).

![Diagram of UK e-Science Programme Structure and Funding](image)

**Figure 14: Structure and Funding for UK e-Science Programme (Hey and Trefethen, 2002)**
While the UK expects investment from industry, in Australia the initiatives were largely funded through government alone (Alberta’s Cyber Infrastructure taskforce, 2006). Many authors indicate that to implement and sustain e-Research initiatives requires considerable investment and support from government.

AERIC (The Australian e-Research Infrastructure Council, 2006) indicated that there is tension between encouraging collaboration in the investment in and use of research infrastructure on the one hand, and competitive funding for research on the other. They further indicate that increasingly, researchers are being asked to address problems that cannot be solved by teams within a single institution or organisation. The research environment now has a larger number of participants, and to achieve credibility, research teams must pool resources and combine intellectual power (AERIC, 2006).

Conversely, the majority of funding for research is distributed on a competitive basis. This is often the case for specific research projects, where funding is usually targeted within a single discipline (AERIC, 2006). While competition itself can encourage innovation, it is at odds with the requirement to pool resources to remain competitive on the global stage (AERIC, 2006).

### 3.5.4 South African Funding Structure for e-Research Activities

Major e-Research activities or infrastructure development, in South Africa have largely been funded by the Department of Science and Technology (DST, 2009). The CSIR’s Meraka Institute has been responsible for the implementation of the Department of Science and Technology’s cyberinfrastructure, which in addition to SANReN (Backbone Network), comprises the Centre for High Performance Computing and the proposed “very large datasets data-storage initiative” (DST, 2009). A clear indication of the total investment made into these infrastructures could not be found.
Having a single Department fund all e-Research activities is perhaps not ideal. As Appelbe (2010) indicated from lessons learned in Australia, it may be necessary to form partnerships in order to fund e-Research activities. As Appelbe (2010) further indicated these partnerships should possibly be comprised of universities, research institutions, provincial, and federal funding. However, the UK programme showed that investments by industry (Hey and Trefethen, 2002) could also assist in the advancement of e-Research. Funding e-Research can therefore be seen to also require collaboration between various stakeholders.

Particularly at the national level, the benefits of collaboration are many, and include cost-sharing across infrastructure investments, the potential to develop new research methods, new cross-discipline approaches, new relationships, and enhanced research outputs (AERIC, 2006).

This leads to the need for greater coordination and oversight, especially where large scale collaboration is being undertaken.

3.5.5 Coordination and Oversight
The Australian Government (2006) in their e-Research Strategy and Implementation Framework report indicated that a committee should be established to provide the strategic direction, drive, engagement and coordination of research groups involved in e-Research. The report further indicated that this committee should be established at a high level and answerable to the Minister for Education, Science and Training and the Minister for Communications, Information Technology and the Arts, and be chaired by the national e-Research Leader.

In Australia, the role of co-ordination of activities is largely the responsibility of AERIC – Australian e-Research Infrastructure Council (AERIC, 2006). AERIC was established by the Department of
Education, Science and Training (DEST) in July 2007 to oversee the Platforms for Collaboration investments and to report to, and be guided by, the then National Collaborative Research Infrastructure Strategy (NCRIS) Committee (NRIC).

The JISC in the UK is largely responsible for co-ordinating e-Research activities for education and research (JISC, 2011). The JISC has been responsible for starting many projects related to e-Research but also for the coordination, oversight and funding of these activities (JISC, 2011).

From the UK and Australian programs, it was shown that there is a need for the coordination of tasks across the full spectrum of e-Research activities. To manage this coordination of undertakings, it can be seen that both the UK and Australia established committee’s or teams to oversee the various activities. This, too, could possibly succeed in the South African context.

As e-Research calls for a fundamental change in the way researchers work, many authors highlight the need for culture change.
### 3.6 Culture

e-Research technologies should be used to underpin research, cultural change in research institutions and amongst researchers, and to provide a platform for sharing collective findings and methodologies between the research, businesses and government sectors (AERIC, 2006). According to the OSI, e-Infrastructure Working Group (2004), the culture of sharing on an international scale and working in cross-cultural global working groups calls for a culture change for researchers and research institutes. According to van Deventer (2011), there is a fundamental change for researchers, who previously worked individually or in small groups, to working in large international teams, where cultural differences at various levels may influence work culture.

In relation to culture, authors highlight the need for leadership to bring about the required change, in order to implement, drive and coordinate e-Research activities.

#### 3.6.1 Leadership

International experience, particularly in the UK, strongly suggests that success in engaging the research community in e-Research is associated with the appointment of a “champion”, a well-respected member of the research community, recognised by government, the research sector and business as the visible face of e-Research (Sargent, 2006). Sargent (2006) and Sargent (2007) highlight that for the e-Research strategy to be effective; it should be championed at the highest levels of the research community, governments and industry.

A champion is further needed to focus the leadership of e-Research by articulating the vision and engaging key stakeholders in taking the strategy forward (Sargent, 2006, Houghton, et al., 2003). This requires a clear focus and coordination of activities across the broad spectrum
of e-Research stakeholders. According to (Sargent, 2006), the implementation strategy for e-Research will require strong leadership as the catalyst to drive the agenda. The appointment of such a person was a critical success factor for the e-Science Initiative in the UK (Sargent, 2006).

The Australian Government (2006) indicated that the e-Research champion should shape the culture and dynamics of the e-Research strategy. They further indicated that the e-Research champion should exist within the operational structure, including within governments, universities, publicly funded research institutions and industry, and should influence these organisations to transform practices and cultural behaviours, and generate enthusiasm for e-Research methodologies.

Furthering the e-Research agenda is as much about people as it is about technology, because e-Research is challenging existing research practices and cultures (Houghton, et al., 2003). There is a need for strong leadership to articulate the vision and engage key players in taking the strategy forward in a coordinated manner (Sargent, 2006). Lord and MacDonald (2003) agree with Sargent (2006) and highlight that a culture change may be required in research environments.

Sargent (2006) recommended in his report that the government should appoint an e-Research Leader to facilitate the e-Research reforms and to champion the e-Research vision. He further recommends that an e-Research Committee be appointed to steer the implementation of the e-Research Strategic Framework, in consultation with the research sector and other stakeholders.
3.6.2 **Technical Expertise and Skills Development**

In Australia, it was found that a deeper and more rapid up-take of e-Research would require a considerable increase in the number of researchers with the confidence, knowledge and technical skills to undertake e-Research, and in the number of IT specialists who could support or participate in e-Research projects with domain-based researchers (Australian Government, 2006).

The NRIC (2006) found that investment would be needed in the expertise and capabilities required to address the many technical challenges to be solved in developing enabling platforms and applying them effectively to the task of producing more collaborative and better research. In addition to the expertise required in the development and implementation phases, another significant issue is the skill sets required to support the researchers (NRIC, 2006). Not every “researcher can be a programmer, or a digital librarian, in addition to meeting the extensive demands of their own particular professional discipline” (NRIC, 2006: 53). The NRIC (2006) found, therefore, that it is important to develop and reward the new and emerging occupations that are able to provide the necessary technical expertise.

The NRIC (2006) advised that in order to have the best chances of ensuring the highest quality content of research outputs at the earliest stage of the research process, it may also be necessary to incorporate information management professionals into research teams. They further indicate that such professionals may assist researchers with their information management needs and also develop guidelines, tools etc, for particular disciplines or particular needs.

Sargent (2007) and O’Brien (2005) highlighted the need for skills development. Sargent (2007) indicated in the e-Research strategy for Australia that three particular skills groups would need to be established or strengthened in order to successfully implement e-Research. The three skills groups identified by Sargent (2007) are:
- Researchers who need to be confident and proficient in e-Research
- Professionals who can translate between research domains and the necessary ICT specialist knowledge, to assist researchers to use advanced ICT and information management methodologies; and
- experts in ICT and information management who can apply their high level knowledge and ability to particular research projects and to the research, development and deployment of e-Research enabling tools, services and methodologies.

In the UK, the JISC also saw a need for skills development and established the JISC NetSkills project, which provides staff development services (JISC, 2011). Continual development of staff skills is vital to universities and colleges seeking to fully embrace the challenges and opportunities arising from the adoption of new and innovative technology (JISC, 2011).

The purpose of skills development activities within the e-Research Strategic Framework, would be to provide capable e-Research practitioners for universities and industry, to train experts to assist researchers to use advanced ICT systems and tools in their research and to create a pool of professionals who can design, operate and maintain advanced ICT systems and platforms on which e-Research will depend (Australian Government, 2006).

From the literature reviewed in sections 3.5 and 3.6, authors have highlighted that various activities, such as DST funding, co-ordination of activities and skills development would be necessary to implement e-Research nationally and in individual institutions. These components indicated can thus possibly form part of the components making up the framework for e-Research as shown in Figure 15 below.
It was established that e-Research is composed of many differing components and also activities are required, they further highlight that collaboration on all levels is essential to bringing about e-Research.

3.7 Summary

The chapter provided an overview of the various components making up the e-Research paradigm as described by various authors. In terms of systems, various components are explained, such as grid technologies, digital repositories and networks.

The funding structures of the UK and Australia e-Research initiatives are described and intangible aspects such as leadership and coordination are highlighted as being important to the successful implementation of an e-Research strategy. The need for skills development is emphasised towards the end of the chapter.

The literature review led to the identification of elements/components that should be included in a framework for the implementation of e-Research in South Africa so that institutions conducting research could also plan their own participation in this new research paradigm.

The literature review further highlighted that most of the components identified in the Australian and UK e-Research programs were funded by government, or through government established institutions responsible for funding and overseeing e-Research activities in those
countries. This has led to an early identification of the need to conduct further research on what format for funding would be applicable to a developing country such as South Africa.

The next chapter, chapter 4, explains the research design and methodology used for the study. Data collection and analysis methods for this study are explained in this chapter.
CHAPTER 4 : RESEARCH METHODOLOGY

4.1 Introduction

The purpose of research can be seen as an investigation into specific issues in business (Coldwell and Herbst, 2004), as well as to expand what was found in the literature to include local findings, in order to build a deeper understanding thereof (Coldwell and Herbst, 2004). This research set out to establish a more holistic understanding of the various components that make up the e-Research paradigm. It is hoped that a deeper understanding of the holistic picture and the various components that make up the picture will assist organisations in their strategy development and implementation of e-Research initiatives.

Although the literature review reported here in Chapter 3, is extensive it was seen as inappropriate to use only literature when designing a framework that will involve such a wide array of stakeholders. It was deemed necessary to use qualitative methods to gain feedback on and further insight into the design of an e-Research framework for South Africa. In particular it was decided to use a focus group discussion and semi-structured interviews to do so. Both of these methodologies as well as the reasoning behind doing so are described in more detail in the sections that follow.

4.2 Research Design Principles

Bryman and Bell (2007) indicate that the purpose of research design is to provide a road map of the whole research project. It should include clear guidelines and procedures with regard to what one intends to do and when. Cooper and Schindler (2001) in Coldwell and Herbst (2004) indicate that while there are differing views on what a research design is, there are underlying principles which are common to the various research design definitions. These principles as defined by Cooper
and Schindler (2001) in Coldwell and Herbst (2004) as well as Bryman and Bell (2007) could be consolidated to be as follows:

- The design is an activity, a time based plan;
- The design is based on the research question. The question that was studied further in this research was to determine which components make up the e-Research paradigm and how or at what level (national or institutional) these components could possibly be funded;
- The design guides the selection of sources and types of information;
- The design is a framework for specifying the relationship among the study’s variables and therefore a framework was designed and will be tested with knowledgeable experts in a focus group discussion;
- The design provides a framework of evidence;
- The design outlines procedures for every research activity which ensures that the research is reliable, replicable and valid.

Bryman and Bell (2007) also indicate that it is essential to link a research design, such as experimental research, comparative research or cross-sectional research designs to a research method, such as qualitative or quantitative research, when planning a research project. As was indicated it was found that qualitative research would be most appropriate approach and method in this instance.

### 4.3 Research Approach and Method

In understanding research and the need for methods or approaches to research, De Vos, Strydom, Fouché & Delport (2005), indicate that irrespective of what one wants to find out, or what facts one wishes to acquire, there is a process involved, a process of enquiry. They therefore indicate that a research methodology refers to the way in which to solve problems, i.e. the research process. De Vos, et al., (2005) indicate that there are two well-known and recognised approaches to the research process, being the qualitative and quantitative approaches.
Myers (2009) indicated that qualitative research methods were developed in the social sciences to enable researchers to study social and cultural phenomena, while quantitative methods of research are described by Charlesworth, Lawton, Lewis, Martin & Taylor (2003) as having been derived from scientific practise and is dependent on the use of scientific method, including statistical analysis.

Coldwell and Herbst (2004) indicate that information is considered to be qualitative in nature if it cannot be analysed by mathematical techniques, while when referring to methods of research Bryman and Bell (2007), and indicate that qualitative research may be construed as a research strategy that usually emphasises words rather than quantification in the collection and analysis of data. Silbergh (2001) indicated on the other hand, that qualitative research methods should not be seen as mere description. Silbergh (2001) goes on to state that though numerical procedures are not essentially involved, logical testing and argument are just as important as in more quantitative methods.

Bryman and Bell (2007) infer that the quantitative researcher is essentially preoccupied with measurement, causality, generalisation and replication. The qualitative researcher on the other hand, is described as being preoccupied with seeing through the eyes of the people being studied, with description and emphasis on context and flexibility, and limited structure, concepts and theory being grounded in data.

Based on the differences stated between qualitative and quantitative research, a qualitative approach has been selected as the approach to be taken for this research study. The information being analysed cannot be undertaken in a numerical way, and the focus of the study will predominantly emphasise an inductive approach to the relationship between theory and research, in which the emphasis is placed on the
generation of theories. For this reason, a qualitative approach rather than a mixed approach or quantitative approach, will be taken to the research.

The review of literature existing on the topic provided a good foundation for understanding the basics of e-Research. The research can be partially described as an internet based content analysis.

**Objective 1:**
To investigate and identify the components of the e-Research and to determine sources of funding for these components, as identified in the United Kingdom (UK) e-Science, and the Australian e-Research programs.

**Objective 2:**
To investigate and identify the components of the e-Research, as implemented in South Africa.

For the first two objectives of the research project, an extensive review of existing literature was undertaken. (As the essence of e-Research is about internet based open access to information, much of the information relating to the various e-Research programs could easily be found on the internet).

**Objective 3:**
To elaborate on and extend the framework of components of e-Research for South Africa and make recommendations for those components that should be funded centrally by government, and those that should be funded by the individual institutions.

Based on the findings from the literature, the SARIS framework of e-Research components was extended and elaborated on.

**Objective 4:**
To calibrate the framework and suggested funding division (government vs. organisation level) with at least one large research institution.
A focus group discussion was selected specifically to validate the framework and to enhance recommendations. Bryman and Bell (2007) indicate that the focus group technique is a method of interviewing that involves more than one individual. Usually at least three interviewees are given the opportunity to discuss an issue or issues. They further indicate that the purpose of a focus group is typically to emphasise a specific theme or topic that is explored in depth. Bryman and Bell (2007) also indicate that the focus group practitioner is invariably interested in the ways in which individuals discuss a certain issue as members of a group, rather than simply as individuals.

Because the SARIS group did not include South Africans currently involved in e-Research infrastructure development, it was decided to use a second quantitative method, semi-structured interviewing, to gain feedback from senior staff already appointed to address specific aspects of the proposed framework. A semi-structured interview, according to Bryman and Bell (2007), typically refers to a context in which the interviewer has a series of questions that are in a general form of an interview schedule, but is able to vary the sequence of questions. The questions are somewhat more general in their frame of reference than those typically found in a structured interview (Bryman and Bell, 2007). This form of interview also provides the interviewer latitude to ask further questions in response to what are seen as significant answers (Bryman and Bell, 2007). This method was used to gain further insight into how the framework could be adapted.

4.4 Data Collection Methods

Firstly, available literature on the topic was reviewed. The main purpose of reviewing the available literature, according to Bryman and Bell (2007), is to determine what is already known about the area of interest. A comprehensive literature review was undertaken.
After completion of the literature review, clearer paths for extending the existing e-Research framework become evident. When a preliminary design for the framework of e-Research components was established, a focus group was used to verify the findings and improve the validity of the research. To further supplement the focus group’s findings, semi-structured interviews were conducted. An iterative approach was used as new information came to light.

### 4.5 Population and Sample

The population refers to the complete ‘group of people, items or units under investigation’ (Coldwell and Herbst, 2004).

Eisenhardt (1989: 537) supports the view that random sampling for a qualitative study is not preferable. The emphasis should rather be to choose participants who are likely to replicate or extend the emergent theory. To meet the objectives of this study, purposive sampling will be employed. Purposive sampling is described by Leedy and Ormrod (2001) to be a sample of people or units selected for a particular purpose.

The members of the focus group were identified with the assistance of the original SARIS group. The members are e-Research stakeholders and included selected members from universities, libraries, research institutes, government and members responsible for high performance computing infrastructure in the country. This should be regarded as purposive sampling rather than ensuring representivity.

As e-Research is a relatively new concept, selecting a broad sample, where people comprising the group may not have the necessary background, experience or information about the paradigm, may skew results. The implementation framework would be better tested against a knowledgeable audience who could provide valuable feedback on the framework. Purposive sampling was used for determining the focus
group participants and well as the participants for the semi-structured interviews.

4.6 Data Analysis methods

When all information has been collected for the content analysis, an analysis will be done on the data. The content analysis aims to quantify content in terms of predetermined categories, and in a systematic and replicable way (Bryman and Bell, 2007). In order to determine the categories, the task of coding will be undertaken. Coldwell and Herbst (2004) indicate that coding one’s data refers to using a code or symbol, usually numerical, to represent responses to survey questions. They further indicate that it is a procedure by which data is organised and categorised into a number of classes. This is done to reduce the large volume of data into smaller groups that contain critical information needed for analysis.

A comprehensive content analysis of the data will be undertaken. Bryman and Bell (2007), however, indicate that one distinct disadvantage of content analysis is the fact that it can be very labour-intensive, and can only be as good as the documents are, on which the researcher is working. Advantages of content analysis include its economy in terms of both time and money, while also allowing for the correction of errors to small areas rather than having to redo the entire process (Babbie, 2008).

4.7 Validity and Reliability

Validity is not possible without reliability; however reliability is possible without validity (Coldwell and Herbst, 2004). Validity and reliability, according to Neuman (2006) are usually complementary concepts, although, in some situations, they contradict each other. Sometimes, as validity increases, reliability is more difficult to attain, and vice versa.
Validity, according to Neuman (2006), suggests truthfulness. Gravetter and Forzano (2003) in De Vos, et al. (2005), refer to validity as being the extent to which an empirical measure accurately reflects the concept it is intended to measure. On the other hand, reliability, according to Neuman (2006), means dependability and consistency. "Consistency is the hallmark of reliability." (Coldwell and Herbst, 2004: 17). Perfect reliability is virtually impossible to achieve, according to Neuman (2006). Rather, it is an ideal all researchers should strive for.

To improve the validity and reliability, a focus group was selected. To further improve the validity and reliability of the findings of the focus group, semi-structured interviews were conducted to further improve the validity and reliability of the findings of the focus group. The use of existing literature in the research study was also aimed at improving the construct validity of the research.

4.8 Summary

Chapter 4 explained the procedures and techniques that were employed to explore the research objectives. The chapter discussed and justified the research design, the data collection procedures, data analysis techniques and population and sample used in the research study.
CHAPTER 5 : RESEARCH RESULTS & DISCUSSION

5.1 Introduction

As discussed in Chapter 1 of this report, this study set out to gain a better understanding of the various components making up the e-Research paradigm, and to investigate how these components are funded. This study further aimed to provide a framework for institutions to use when implementing an e-Research strategy, to guide them on which components they should be implementing.

For the semi-structured interviews, key role players were interviewed. The key role players interviewed are as follows:

- Dr Daniel Adams, Chief Director: Emerging Research Areas & Infrastructure, Department of Science & Technology;
- Dr Colin Wright, Manager: Cyberinfrastructure, The CSIR’s Meraka Institute.
- Dr Bruce Becker, Co-ordinator for SAGrid, The CSIR’s Meraka Institute.
- Mr Roy Page-Shipp, Retired SARIS project team leader and part time lecturer in Information and Knowledge Management at GIBS.

Contributions made by these role players and participants of the focus group will be reflected in this chapter anonymously. This chapter reports the results of the study on a per objective basis and indicates the findings from the focus group discussion and semi-structured interviews undertaken.
5.2 Findings for Research Objective 1: UK and Australian Components of e-Research & Funding

The results for Objective 1 were reported extensively in Chapter 3 and after the qualitative research it was confirmed that a South African Model would look similar to that of the UK and Australia.

**Objective 1:**
To investigate and identify the components of the e-Research and to determine sources of funding for these components, as identified in the United Kingdom (UK) e-Science, and the Australian e-Research programs.

The following section reports the findings for the above objective and indicates both the components of e-Research identified as well as the various sources of funding for both the UK and Australian e-Research programs.

5.2.1 *Components of e-Research*

It was found from the literature review (Chapter 3) done on the UK and Australia e-Research programmes, that there is largely a consensus on the basic components of e-Research. It was found that both the UK and Australian programs built their e-Research around a grid/cloud architecture or distributed computing architecture (see Section 3.1), comprised of high speed networks, high performance computing and digital repositories. Three of the interviewees confirmed that South Africa would need to do the same, and in most respects have already begun implementation of these “cyberinfrastructure” components.

Both programs (UK and Australia) indicated that middleware (see section 3.1.4) would be essential to bring these various components together and to allow access to the underlying architecture. One of the interviewees confirmed that grid middleware would be especially important to bring together the functionalities of the underlying
infrastructure. All interviewees did not dispute the need for middleware and services.

It was further found that the UK, through the JISC (2011), has established a number of programmes to address the researcher’s needs in terms of tools and applications. These include tools that address visualisation, access, authorisation and authentication, access to information, communication & collaboration, digital curation and preservation, e-learning and project-specific tools. In Australia, many of these needs are also being addressed through national programmes. In the focus group discussion, it was highlighted that the development of these tools is not receiving much attention in South Africa.

The Australian programme highlighted the need for a body/committee such as AERIC. This body would be responsible for the co-ordination of all research activities across the overall investment. The focus group highlighted that such a body would be necessary in South Africa, but highlighted that it should be comprised of multi-disciplinary members.

5.2.2 Sources of Funding for e-Research

In Australia, from the literature (see section 3.5.2), it was found that the government funded many of the e-Research projects, through the NRIC (2006). From the information found on the funding of e-Research activities it was found that the NCRIS’s policies for funding are such that major infrastructure is developed on a collaborative, national, non-exclusive basis. Infrastructure funded through NCRIS would serve the research and innovation system broadly, not just the host/funded institutions. Funding and eligibility rules established by the Australian Government encourage collaboration and co-investment. It was further found that the NCRIS would not fund institutional level (or even small-scale collaborative) infrastructure. In the Australian programme for e-Research, AERIC (2006) indicated that there is tension between
encouraging collaboration in the investment in and use of research infrastructure on the one hand, and competitive funding for research on the other.

This way of funding e-Research, however, places a large demand on resources provided by government alone. In countries like South Africa, where funding is scarce, government funding alone may not be sufficient to implement a sustainable e-Research framework. The focus group and three of the interviewees indicated that while DST currently funds many of the e-Research activities in South Africa, for this model to be sustainable, external funding/private funding would be required.

In the UK, from the literature, section 3.5.3, it was found that while a large proportion of the funding for e-Research came from government, there were a number of partnerships established, especially between research councils and education departments to fund the infrastructure. As is common in Department of Trade and Industry (DTI) programmes, the DTI contribution required a matching contribution from industry (Hey and Trefethen, 2002). The funding in the UK was funnelled to a central body – JISC. JISC was then responsible for distribution of the funds, from the funding partnerships, into e-Research projects and activities as are required.

The UK model of funding can be considered to be more sustainable over the long term than the Australian model is, especially if being implemented in a developing country. However, this may call for the review of national funding policies. The focus group agreed that in the longer term, the UK model for funding would better suit the needs of South Africa.
5.3 Findings for Research Objective 2: South African Components of e-Research & Funding

Literature, describing the components to be considered for Objective 2, was reported extensively in Chapter 3. After the qualitative research, it is possible to report that many of the components for infrastructure have either already been implemented or are in the process of being implemented.

<table>
<thead>
<tr>
<th>Objective 2:</th>
</tr>
</thead>
<tbody>
<tr>
<td>To investigate and identify the components of the e-Research, as implemented in South Africa.</td>
</tr>
</tbody>
</table>

It was established that South Africa has focused on building a “cyberinfrastructure”, which is inclusive of grid architecture (SAGrid), high speed backbone network (SANReN) and High Performance Computing (HPC). The foundation has further been laid for a digital data repository, known as the “Very Large Database”, although it is still to be implemented. This was confirmed by two of the interviewees. It was further found (see section 3.2.1) that there are nine open access repositories at several academic institutions.

In the focus group discussion, it came to light that programmes for implementing specific components of e-Research were being undertaken, for example, the Academy of Science for South Africa, has programmes focusing on improving visibility and open access to journals, as well as skills development.

The focus group and one of the interviewees strongly indicated the need for a strategic view of the e-Research developments currently being undertaken. It was highlighted that there are possibly many developments taking place, but other institutions may not be aware of them, and could possibly be duplicating work. This emphasises the need for a body such as Australia’s NRIC, to oversee all e-Research
activities, and the UK’s, establishment of JISC, to ensure cost-effective and strategic oversight of the activities taking place. This was confirmed as being important by the focus group. One of the interviewees also indicated that strategic oversight would be necessary and that the members making up the oversight committee could participate in policy development.

The South African “cyberinfrastructure” has largely been funded by the Department of Science and Technology project managed by the Meraka Institute, a division of the CSIR. This was confirmed by three of the interviewees. In terms of funding there was found to be little indication that government policies promoted co-investment and collaboration, such as was seen in the Australian program. This was highlighted by the focus group and one of the interviewees and was indicated as possibly being a barrier to the implementation of e-Research. The focus group and two of the interviewees suggested that collaboration on all levels is necessary for e-Research to be truly successful.

5.4 Findings for Research Objective 3 & 4: A calibrated framework of components for e-Research

Literature, describing the components to be considered for the extended framework, was reported extensively in Chapter 3. After the qualitative research, it is possible to report that many of the components indicated in the framework, should be included, with some modifications.

| Objective 3: |
| To elaborate on and extend the framework of components of e-Research for South Africa and make recommendations for those components that should be funded centrally by government, and those that should be funded by the individual institutions. |

| Objective 4: |
| To calibrate the framework and suggested funding division (government vs. organisation level) with at least one large research institution. |
From the review of literature in Chapter 3, a preliminary framework for components of e-Research was drafted (Shown in Appendix 1: Figure 27). This framework was presented to the focus group and interviewees for further discussion.

5.4.1 **Components of e-Research**

The focus group voiced a concern that the framework did not clearly show how the various components making up e-Research flow into one another or co-exist. This was also indicated by Interviewee A, in an interview, where it was highlighted that a framework of this magnitude was difficult to show in two dimensions. It was suggested that research currently being undertaken shows that there are possibly five dimensions to e-Research. It was further highlighted that two important considerations should be that of scale and complexity, in that some researcher communities may have large scale data but very little complexity involved in the data, while others may be opposite. Interviewee A indicated that as a two dimensional depiction of e-Research, the framework indicated many of the components necessary for e-Research. Two of the interviewees, as well as most of the participants of the focus group, found that the preliminary framework contained most of the components making up the e-Research paradigm. Interviewee C stated that the framework was comprehensive and indicated the correct layered structure. Interviewee C opined that the framework showed what the e-Research landscape should be.

The various layers making up the components of the e-Research framework are discussed in the sections below.
5.4.1.1 e-Research Infrastructure or “Cyberinfrastructure”

Components

The focus group and the interviewees confirmed the components identified (see Figure 10 in section 3.1). However, Interviewee D indicated that some institutions might not require high performance computing but would require computing infrastructure to process and capture information. The focus group and two of the interviewees indicated that due to the small number of institutions requiring network connectivity there might not be the need for both provincial and inter-institutional connectivity, but that it could be one or the other. Interviewee B suggested that the wording of provincial networks be changed to regional networks, as this would indicate an area, but not limited to a province.

The focus group discussed digital repositories and high performance computing and indicated that they were composite of two aspects, the infrastructure needed for these repositories or computing, and the services required to be able to store information. Interviewee C indicated that there should be national digital repositories, as well as institutional digital repositories. For these reasons, the infrastructure layer has been changed to be as follows:

<table>
<thead>
<tr>
<th>High Performance Computing Infrastructure</th>
<th>Computing Infrastructure</th>
<th>Data Storage Infrastructure/Repositories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regional Networks &amp; Inter-Institutional Networks</td>
<td></td>
<td></td>
</tr>
<tr>
<td>National Backbone Network</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 16: e-Research Infrastructure or “cyberinfrastructure”

Built on top of the infrastructure layer, a middleware and services layer is indicated. The middleware and services layer is discussed in the section below.
5.4.1.2 Middleware & Services

The components identified (see Figure 11 in 3.2) relating to middleware were not highlighted by the focus group or any of the interviewees as being incorrect. In an interview with Interviewee D, the distinction between information and data middleware and knowledge management middleware was discussed. It was agreed by Interviewee D that there should be the distinction between the two, to bring to the fore the need for knowledge management over and above data and information management. Interviewee A agreed that grid middleware should cover the entire infrastructure layer. For these reasons, the middleware layer remained unchanged as indicated in the figure below.

Building on the middleware and services layer is an applications layer. The applications layer is discussed in the section below.

5.4.2 e-Research Applications

The components identified (see Figure 11 in section 3.2 and Figure 12 in section 3.3) were indicated as being tools and applications. The focus group stated that some of these components should in fact be classified as products and services, and separate to this would be applications. Interviewee B agreed with this distinction. Examples were given such as visualisation and project specific tools, which should rather be classified as applications, while open access to information should be classified as products and services. The focus group further discussed high performance computing and indicated that high performance computing spread across both infrastructure and applications. For these reasons, an applications layer was brought in
and some of the components previously indicated as being tools and applications were brought into this layer, as shown below.

<table>
<thead>
<tr>
<th>High Performance Computing Applications</th>
<th>E-Learning/ Digital Scholarship Tools &amp; Applications</th>
<th>Project Specific Tools &amp; Applications</th>
<th>Visualisation</th>
</tr>
</thead>
</table>

**Figure 18: e-Research Applications**

Building on the applications layer is a products and services layer. The products and services layer is discussed in the section below.

**5.4.2.1 Products and Services**

A products and services layer would then be built on top of this applications layer, as indicated by the focus group. Interviewee B indicated that a further service required, which had not been included to date, would be “Digitisation”. Interviewee D agreed with this. Interviewee D also indicated that researchers require products and services which give them access to remote instrumentation, such as telescopes and electron microscopes, which may be located on different continents. Access to information was found by the focus group to be sufficient, but that there should be two separate components; one for commercial information, and one for open access to information. This is in line with the findings of Houghton, et al., (2003), that open access digital repositories, operating in parallel with existing commercial publishing mechanisms, can provide a major opportunity to develop a sustainable information infrastructure for both traditional and emerging modes of knowledge production.

The focus group further indicated that products for data storage and services relating to this should be included, which would be separate to the data storage infrastructure. Interviewee A suggested the need for an application database component, which would be a database recording all e-Research development activities, specifically indicating when a component was developed, then who owned it, who was
responsible for running it and so on. This database would assist organisations that wish to implement a component in ascertaining who they could approach to assist them in gaining access to the resources or advice on how they could implement the component in their organisation. Interviewee D agreed with this. For these reasons, the products and services layer was adapted and changed to be as follows:

<table>
<thead>
<tr>
<th>Communication &amp; Collaboration Services</th>
<th>Digital Curation &amp; Preservation Services</th>
<th>Access to Licensed or Commercial Data &amp; Information</th>
<th>Remote Instrumentation</th>
<th>Open Access Data &amp; Information &amp; Services &amp; Products</th>
<th>Primary Data Sharing</th>
<th>Digitisation</th>
<th>e-Research Information Database</th>
<th>Large Scale Data Storage Services</th>
</tr>
</thead>
</table>

![Figure 19: e-Research Products and Services](image)

It is very possible that this layer specifically would be expanded as more stakeholders acquire the opportunity to evaluate the framework. Expanding on the products and services layer are further tools required for accessing the services, products and infrastructure. This layer is discussed in the section below.

### 5.4.2.2 Access & Mobile/Remote Connectivity Components

The focus group and two of the interviewees indicated the need for mobile connectivity to this infrastructure and services. The mobile or remote connectivity is to allow researchers to connect to the infrastructure regardless of where they are located in the world. The focus group also indicated that access, authentication and authorisation (as discussed in section 3.3.4 and shown in Figure 12) as a service would be cross cutting; thus would affect all products, services, middleware and infrastructure. Another such service would be that of Grid/Cloud Access to products and services which was highlighted by two of the interviewees as being necessary to access all infrastructure and services. This access allows the user to make
use of services and infrastructure seamlessly, as well as providing on-demand resources as required. Mobile or Remote Connectivity was not highlighted in the literature, as authors did not specifically highlight the need for this. However, with rapidly changing technologies, the need for mobile connectivity is becoming increasingly important, as indicated by the interviewees. For these reasons a mobile connectivity model was included and the Access model was modified as shown below.

<table>
<thead>
<tr>
<th>Mobile/Remote Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grid/Cloud Access to Products, Services, Applications &amp; Infrastructure</td>
</tr>
<tr>
<td>Access, Authentication &amp; Authorisation</td>
</tr>
</tbody>
</table>

*Figure 20: Access Modules for e-Research*

Further services required to bring about e-Research, such as training are discussed in the section below.

### 5.4.2.3 Quality Assurance, Training & Skills Development Components

As indicated by authors in section 3.2.2, the focus group confirmed that quality assurance would be essential to ensure the quality of the information made accessible to researchers. Without quality control South African research output would not be globally acceptable.

The focus group highlighted that user training services would be essential to provide the necessary skills development in utilising the underlying products, services, applications and infrastructure correctly and efficiently. This should be differentiated from just information literacy or computer skills development. The focus group also indicated that skills development (see Figure 15, section 3.6.2) should be depicted as being overlapping and modular as it is unlikely that an specific individual would find it useful to have the skills required at all levels of the framework. Two of the interviewees agreed with this. To
bring about this level of change, a culture change should be considered. Interviewee A further indicated that a training infrastructure would be required, as training should ideally not occur on live production systems. He indicated that by using cloud computing a dynamic management model could be used to separate tools and infrastructure for training purposes. Interviewee D agreed with this need for training infrastructure. For these reasons, the skills development component was modified, training infrastructure and user training services and as well as quality assurance services added, as shown below.

<table>
<thead>
<tr>
<th>Quality Assurance &amp; User Training Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Infrastructure</td>
</tr>
<tr>
<td>Skills Development</td>
</tr>
<tr>
<td>Skills Development</td>
</tr>
<tr>
<td>Skills Development</td>
</tr>
<tr>
<td>Skills Development</td>
</tr>
<tr>
<td>Skills Development</td>
</tr>
</tbody>
</table>

**Figure 21: User Training, Skills Development & Quality Assurance Components**

The focus Group indicated the need for an access layer which would allow collaborating researchers to group together all the services, infrastructure and applications that they would need to access and use. The focus group identified this ‘access layer’ as being known as the Virtual Research Environment (VRE). Interviewee A emphasised that the VRE would be an essential component. Although VREs were not investigated during the desk research phase of this investigation, it was established that much literature is available on the subject and therefore the layer was added to the model.
5.4.2.4 Users, Researchers and the Virtual Research Environment

The user component indicated on the framework was found by the focus group to be inadequate. It did not clearly indicate that this framework/infrastructure is meant for researchers. They highlighted that the researcher should be made more prominent in the framework, so that if researchers look at the diagram, they would be able to immediately identify where they would fit into the picture. The focus group indicated that e-Research calls for collaboration and as such should be indicated as being e-Researcher communities as e-Research highlights the need for collaboration between and amongst researchers. Interviewee B agreed with this, but Interviewee D argued that certain types of research could still not be undertaken in a collaborative way but rather as individuals. In addition, Interviewee A indicated that developers as well as support staff would also be using this infrastructure, and as such, should be included in the user community.

Interviewee C indicated that “buy-in” from the users would be necessary. For these reasons, the users indicated were expanded to be more inclusive and a VRE (Virtual Research Environment) was added to the framework, as shown below.

![Figure 22: Users and the Virtual Research Environment for e-Research](image)

Beyond physical products, services and applications, oversight and co-ordination are indicated as being important to the implementation of e-Research. These components are discussed in the sections below.
5.4.2.5 **Strategic Oversight, Leadership & Co-ordination Components**

An oversight and leadership committee as identified in Figure 15 and sections 3.5.5 and 3.6.1 was found by the focus group to be necessary. The researcher recognized wariness amongst interviewees as well as the focus group that researchers would be reluctant to have to report to a central committee. Interviewee D indicated that leadership and an e-Research champion or champions driving the cause would be necessary to bring about the change required. Interviewee C indicated that oversight committee members should have a strategic viewpoint and that they could assist agreed that this committee should be made up of multi-disciplinary members, who should be responsible for overseeing the e-Research paradigm holistically.

Interviewee A indicated that a reflection group would be necessary to determine the long term usage of the infrastructure. Interviewee A indicated that this strategic function could be incorporated into the oversight and leadership committee; however, their roles and responsibilities would need to be clearly defined. Interviewee D highlighted that this committee would possibly need to establish working groups to address certain developments or issues arising. Interviewee C agreed with this, and suggested that the working groups would be responsible for the actual research and scoping of smaller parts of the overall e-Research paradigm.

The components identified (see Figure 15 and section 3.5.5) indicated that co-ordination of activities should take place. The focus group and two of the interviewees agreed with this. The focus group indicated that this team of people should be comprised of operational members from various disciplines, thus be multidisciplinary. They also indicated that these members should have the necessary knowledge, expertise and standing within their individual communities, so that they would be able to ensure collaboration amongst all stakeholders. An indication
was given by the focus group that this team should report to the oversight and leadership committee.

For these reasons, the oversight and leadership committee was revised to be shown as being multidisciplinary and strategic, while the co-ordination of activities team remained largely unchanged as shown below.

![Multi-disciplinary Strategic Oversight & Leadership Committee](image)

![Co-Ordination of Activities Management Team](image)

**Figure 23: Committee and Management Team for e-Research**

Multi-disciplinary committees call for collaboration between the various committee members. The collaboration component is therefore discussed in the section below.

**5.4.2.6 Collaboration Component**

Collaboration as a key and necessary driver for e-Research was highlighted by the focus group, and three of the interviewees. From the literature (see section 3.3.1, it was stressed that collaboration tools would be essential to enable researchers. Through the interviews and focus group discussions, it became clear that institutions function in a ‘silo’ like manner, with very little cross institution collaboration.

For the sake of developing the knowledge economy, a national scale collaboration infrastructure, from which all institutions could benefit, should be further investigated. Interviewee C indicated that this would include collaboration in the way researchers work, but also in the development of the components identified. Two of the interviewees highlighted the importance of collaboration throughout the framework. Collaboration, they indicated, should be extended to include the funding for the various components, as often the development is of such a scale that more than one institution may need to co-fund such
development and as such should collaborate. Interviewee C again emphasised that e-Research infrastructure and systems should provide an enabling environment for researchers. For this reason collaboration has been modified as follows:

![Figure 24: Collaboration in e-Research Activities](image)

Having identified the components of e-Research, funding these various components was then addressed briefly. The results from that discussion are reflected in the next section.

### 5.4.3 Funding of e-Research Components

Interviewee A and the focus group indicated that funding and governance should be depicted as being top-down and that there should not be funding and governance coming into the co-ordination of activities management team without transparency to all.

The focus group and two of the interviewees raised concerns related to the funding for such a large and disparate architecture. A debate arose in the focus group around how the funding models in the UK and Australia were being implemented and the success thereof. Consensus was reached in the focus group, that the UK funding model would possibly work best in South Africa, however it was said that this would possibly be a long term goal. Interviewee A, in turn, indicated that funding an infrastructure like e-Research centrally would not be possible. He is of the opinion that each component should be separately funded. He agreed, though, that it could cause duplication in development activities.

The focus group indicated that to ensure collaboration and also to avoid duplication of activities, funding would need to be consolidated
in some way. It was agreed that individual institutions should seek their own funding but that the oversight committee should make it possible to establish whether a specific component had already been developed (or could perhaps be under development) before individuals would seek funding. The focus group and two of the Interviewees indicated that both public and private funding would be necessary to fund the development of the entire infrastructure. Interviewee A indicated that only if private funding were used could true innovation occur in the development of the e-Research Infrastructure. Interviewee C agreed with this, and indicated that the products and services layer of the framework would be where private funding would most likely be required.

Interviewee C indicated that components developed on a National Scale, such as the backbone network, would be largely funded by government, but that networks used within institutions should be funded by the institution itself. Interviewee C indicated that any components required by any institution would have to be funded by that institution itself, so for example, gaining access to visualisation applications would be at the institution’s expense.

It was found that there could not be a clear divide between what should be funded by government and what should be funded by individual institutions. Instead, it was found that funding partnerships would be required. Interviewee C agreed with this. These partnerships could be between different government departments, between different institutions or between government and institutions. Interviewee C indicated that within government, partnerships/collaboration would in all likelihood be required between, amongst others, the Department of Science and Technology, Department of Higher Education, Department of Public Enterprises, and possibly the Department of Trade and Industry, in order to bring about e-Research.
An example of a partnership was given by Interviewee C, indicating that the establishment of the Centre for High Performance Computing was funded by government (70%) and by industry (30%). What appears to be evident is that collaboration is required on all levels from policy development and funding to ensure the development of the various components. For these reasons the funding layer in the framework has been adapted to include a policy development layer incorporating governance, as well as indicating both public and private funding.

![Figure 25: Funding of e-Research Components](image)

Funding of e-Research needs to be studied further and possibly starting from a policy development perspective. As it was noted that a clear divide between governmental and institutional funding cannot be drawn, it can be concluded that collaboration with regards to funding will be essential.

The results presented in the preceding sections of this chapter indicated that the basic components making up the e-Research paradigm would be as indicated in Figure 16 - Figure 25. The holistic, inclusive framework is provided in the next section.

### 5.5 An e-Research Implementation Framework for South African Organisations

This framework, shown in Figure 26 draws on the information presented in the literature (Chapter 3) and incorporates key research findings (Chapter 5) in an integrative manner.

The Framework (Figure 26) is built up of a variety of layers. An infrastructure layer (coloured in blue) forms the base upon which all other layers are built. This infrastructure layer consists of all the
physical hardware necessary to implement e-Research. On top of the infrastructure layer is the middleware layer (coloured in purple). The middleware layer comprises components necessary to allow applications to interact with the underlying hardware in the infrastructure layer. Building on top of the middleware layer, is the applications layer. The applications layer consists of various applications required by researchers. On top of the applications layer, is a products and services layer that consists of various tools, services and access to a variety of information sources.

The four layers described above can then all be accessed by the users through what is known as a Virtual Research Environment and/or through Grid/Cloud access applications. Through all layers, skills development, training infrastructure and services are required to ensure all users, for example, researchers, librarians and IT specialists can support and/or participate in e-Research.

It is further indicated that at least two governance bodies should possibly be established. The first body (management team) would be responsible for the co-ordination of all activities through the various layers of the framework. It would ensure the cost-effective development, implementation and integration of the various layers. The second body (an oversight committee) would be responsible for general oversight, leadership and providing strategic direction. The oversight committee would have to be multi-disciplinary in its constitution and in all possibility would need to be funded through e-Research funding partnerships. This body would be responsible for establishing work groups to research and develop each of the various components e-Research.

This framework can assist government in driving the e-Research paradigm forward in South Africa. It can further assist institutions in implementing e-Research with their organisation, and aid strategy development, specific to e-Research.
**Figure 26: An e-Research Implementation Framework for South African Organisations**

The diagram illustrates the e-Research Implementation Framework for South African organisations, highlighting the integration of Public Funding and Private Funding through e-Research Funding Partnerships. It also emphasizes the role of the Multi-disciplinary Strategic Oversight & Leadership Committee in guiding the development and governance of e-Research initiatives.

### Virtual Research Environment

- **Mobile/Remote Connectivity**
- **Grid and/or Cloud Access to Products, Services, Applications & Infrastructure**
- **Access, Authentication & Authorisation**
- **Quality Assurance & User Training Services**

### Core Components

- **High Performance Computing Applications**
- **E-Learning/Digital Scholarship Tools & Applications**
- **Project Specific Tools & Applications**
- **Visualisation**

### Support Infrastructure

- **Data and Information Middleware & Services**
- **Knowledge Management Middleware & Services**
- **Collaboration Middleware & Services**
- **Grid and/or Cloud Middleware & Services**

---

**Policy Development & Governance**

**e-Researchers Users**

**e-Researcher Communities Developers Support Staff**

**Public Funding**

**Private Funding**

---

The framework integrates various aspects of e-Research, including skills development, training infrastructure, and collaboration, to support the development of high-performance computing, e-learning, project-specific tools, and visualisation services, as well as addressing infrastructure needs such as data storage and regional networks.
The framework, indicated in Figure 26, could be explained as follows:

- **Infrastructure or cyberinfrastructure layer** refers to the physical infrastructure required to build an enabling environment for e-Research, onto which other layers can be built to enhance the use of the infrastructure. This infrastructure is composite of the following:
  - **National Backbone Network**: A high speed, large bandwidth network for all regional and/or inter-institutional networks to connect to, as well as connecting into Africa and internationally to other education and research networks and grids, such as in Europe, Australia and America.
  - **Regional &/or Inter-Institutional Networks**: A network linking all institutions in a region together or a network linking all institutions together, such as all universities in a region, or linking all research institutions together and then connecting to the national backbone network. This can be expanded to have a network linking all schools or all government departments, which could also be connected to the backbone network.
  - **High Performance Computing Infrastructure**: This infrastructure is necessary to allow researchers to process large volumes of data at high speeds or do complex analysis. A High Performance Computing Infrastructure can “enhance significant research, address grand challenges, and grow computational research” (CHPC, 2010).
  - **Computing Infrastructure**: An institution will not only require access to high performance computing infrastructure but also standard computing infrastructure to manage the e-Research applications, products and services. This also incorporates desktop machines and mobile devices, such as cell phones and iPads to allow the researchers to successfully do their research.
  - **Data Storage Infrastructure/Repositories**: This is where digital content and assets are physically stored and can be searched for and retrieved for later use. Data storage can be done on a
national level as well as being necessary within institutions. (Institutions will need to store and manage their data where it does not make sense to make use of central storage or where security issues come into play.) A repository supports mechanisms to import, export, identify, store and retrieve digital assets (JISC, 2011).

- **Middleware & Services** refers to the communications layer that allows applications to interact across hardware and network environments. The middleware layer is composite of the following components:
  - **Grid Middleware & Services**: provides essential access, communication, accounting, security, trust, and co-ordination services between the (computational and data) resources of the grid and the higher-level services that use them (DSTC, 2004).
  - **Data and Information Middleware & Services**: provides tools and services that enable the indexing, archival, discovery, analysis, integration, management and preservation of large heterogeneous distributed data repositories and digital archives (DSTC, 2004).
  - **Knowledge Management Middleware & Services**: The focus of this middleware is to make use of data and information generated, archived, indexed etc, to mobilise knowledge and aiming towards building a knowledge economy.
  - **Collaboration Middleware & Services**: provides tools and services to support formal and informal, real-time and offline collaborative activities between remotely located researchers, research communities, and resources (dynamic, scalable virtual organizations) (DSTC, 2004).
• **Applications** refer to those specific applications required to use the underlying infrastructure and middleware. Not all institutions will require every one of the applications.
  
  o **High Performance Computing Applications** refers to applications required to use the High Performance computing infrastructure.

  o **E-Learning/Digital Scholarship Tools & Applications**: Tools and Applications to provide an innovative learning experience for learners (Li, *et al.*, 2006). This will incorporate a large variety of various applications to enable an e-learning environment.

  o **Visualisation**: Visualisation can be divided into two broad types or categories, Scientific Visualization and Information Visualization (Charters, 2010). Scientific visualization is concerned with the conversion of numeric data, usually from scientific experiments and simulations to a graphical representation. Information visualization on the other hand is concerned with the conversion of other forms of data, structured and unstructured text, images and video, to an appropriate graphical representation (Charter, 2010). This application is essential for researchers.

  o **Project Specific Tools & Applications**: For specific research projects, certain tools may be required and interfacing with these tools will be necessary, for example, telescopes. Dependant of the type of research being undertaken specific tools will be required.

• **Products and Services** refer to those components that researchers may require, however not all institutions will require all of the products and services.

  o **Communication & Collaboration** refers to those products and services that allow researchers to communicate and collaborate, inter-institutionally, nationally and internationally. Collaboration may involve the sharing of data and access to
content in repositories and increasingly communication and conferencing (Sargent, 2006), including audio conferencing, video conferencing and instant messaging.

- **Digital Curation & Preservation:** This activity refers to the active management of datasets for their scientific and scholarly useful lifetimes, including the promotion of effective and widespread use. Other ePrint archives, for example, theses and research reports, form part of this category (Page-Shipp, et al., 2005).

- **Access to Licensed or Commercial Data and Information:** The pursuit of high quality research requires ready access to the published work and data of other researchers, and the facility to publish one's own findings (Page-Shipp, et al., 2005). If licences are negotiated on a national level, researchers in any institutions will have better access to commercial information.

- **Open Access Data and Information Services & Products** refers to those products, like online publications in formal journals, but where a licensing fee is not required. It also refers to the access to national and institutional repositories. Open access digital repositories, operating in parallel with existing commercial publishing mechanisms, can provide a major opportunity to develop a sustainable information infrastructure for both traditional and emerging modes of knowledge production (Houghton, et al., 2003). Together, they provide the foundation for more effective and efficient access to, and dissemination of, scientific and scholarly information.

- **Remote Instrumentation** refers to those services that allow researcher to remotely control instrumentation and equipment.

- **Primary Data Sharing:** Making primary research data available to other researchers is part of the e-Science paradigm. One needs to store not only the database *per se*, but also sufficient metadata to enable a potential user to find
relevant data and be satisfied with its value and provenance (Page-Shipp, et al., 2005).

- **Digitisation** refers to those products and services required to convert data and information into a digital format, to be further used by researchers.

- **e-Research Information Database** refers to a database of information specifically relating to the development of the various components of e-Research. This database should highlight specific developments being undertaken, who is responsible for them and who to contact to gain access. This database is aimed at institutions that are looking to gain access to or develop their e-Research components, which can then look at the database to determine if, for example, visualisation software has already been developed, then where they can acquire the software. In this manner, duplication of development is avoided, costs are saved, and collaborative relationships are built.

- **Large Scale Data Storage Services** refers to those services, such as indexing, managing and storing of large scale data.

- **Quality Assurance and User Training Services** refers firstly to products or services that ensure the quality of data underlying is of such a quality that it can be used by researchers. The User Training Services refers to those products and services required to enable researchers to use the underlying infrastructure, products, and services.

- **Access, Authentication & Authorisation**: Security technologies must support a variety of user and provider requirements such as authentication, authorization, trust, privacy, policy management, and information assurance, all in a user- and provider friendly framework (ISSR, 2010).

- **Grid and/or Cloud Access** refers to those services that allow Grid and/or Cloud access to the underlying infrastructure, services etc. It allows on-demand access to resources for the user in a seamless way.
• **Users, Access & Mobile/Remote Connectivity** refers to all applications specifically related to the users or being used by users or researchers.
  
  o **Mobile/Remote Connectivity** refers to the tools, products and services that allow researchers to access the products, infrastructure or services mobilely or remotely. This can include cell phones, iPads, and the applications required to run on the products that allow access to the infrastructure from remote locations.
  
  o **Virtual Research Environment (VRE)** refers to the set of online tools, systems, processes interoperating to facilitate or enhance the research process across institutional boundaries. The VRE provides researchers with the tools and services they need to do research of any type efficiently and effectively. The VRE further facilitates collaboration among communities of researchers, often across disciplines and national boundaries.
  
  o **e-Researchers, e-Researcher communities, Users, Developers, Support** refer to all the people who can use, participate in, develop and support the e-Research Framework.

• **Skills Development & Training Infrastructure** refers to the development of skills process and the training infrastructure required to up-skill researchers, support staff, IT specialists, amongst others.
  
  o **Skills Development:** A deeper and more rapid up-take of e-Research would require a considerable increase in the number of researchers with the confidence, knowledge and technical skills to undertake e-Research, and in the number of IT specialists who can support or participate in e-Research projects. Through each of the layers, specific skills will be required for different groups of people. Skills development will be critical in implementing e-Research.
- Training Infrastructure refers to the infrastructure required in order to develop skills. Cloud computing can assist with this as resources can dynamically and in a virtual way be made available for training purposes and then re-assigned on completion of the training.

- Multi-Disciplinary Strategic Oversight & Leadership Committee refers to a committee made up of members with a strategic vision of e-Research from various institutions that are also multi-disciplinary. Its function is to provide the strategic direction, drive, engagement and co-ordination of effort of research groups involved in e-Research. This committee will be responsible for the strategic vision of e-Research on a national scale. The committee should further be responsible for establishing working groups. These working groups would, for example, be responsible for the research and development of one of the components, such as visualisation. The committee should also be responsible for determining the strategic direction of e-Research over the long term. e-Research calls for a culture change among all stakeholders, and to drive this change requires an e-Research ‘champion’ or leader, well-respected members(s) of the research community recognised by government, the research sector and business, as the visible face of e-Research (Sargent, 2006) and to drive the e-Research endeavour. This committee should appoint the necessary leader or leaders.

- Co-Ordination of Activities Management Team: A team to ensure co-ordination of activities, across the various layers of the framework. To be cost effective nationally, all activities should be co-ordinated in their planning, development and implementation. This committee should also be responsible for ensuring collaboration on all levels.
- e-Research Funding Partnerships refers to collaboration between government, industry and institutions to fund e-Research activities, including development and implementation of e-Research components nationally and institutionally. The partnerships will require funding from both public and private sector in order to bring about the e-Research framework.

- Policy Development and Governance highlights the need for policies to be reviewed on all levels and especially in terms of funding, to ensure co-investment and collaboration. Governance refers to the act of governing. It relates to decisions that define expectations, grant power or verify performance. It forms part of the management and leadership processes required for e-Research. This governance should be administered by government.

- Collaboration was found to be a key and necessary driver for e-Research. Collaboration would be required between stakeholders on all fronts. This is inclusive of collaborative funding, collaborative development, and collaboration in conducting research. From a researcher perspective, Katz and Hicks (1997) indicated that the citation of collaborative papers was higher than that of sole-authored papers, and that the impact was higher for collaborative papers where those collaborations were international. Particularly at the national level, the benefits of collaboration are many, and include cost-sharing across infrastructure investments, the potential to develop new research methods, new cross-discipline approaches, new relationships, and enhanced research outputs (AERIC, 2006).
5.6 Summary

Throughout this chapter, the various components making the e-Research framework have been highlighted. Furthermore, the funding of these various components was discussed. In section 5.5, the e-Research implementation framework for South African Organisations was explained, and a brief description of each of the components indicated was given. These findings have built a strong argument in favour of the implementation of this framework in South Africa, both on a national and organisational level.

The next chapter highlights the conclusions of the research, as well as indicating recommendations as to what steps possibly need to be taken in order to implement the framework, indicated in section 5.5.
CHAPTER 6 CONCLUSIONS AND RECOMMENDATIONS

6.1 Introduction

The preceding chapters have examined the components making up the e-Research paradigm, and how these components are being funded. Based on the findings of this research, an e-Research framework was developed and described in Chapter 5, section 5.5. This framework is intended to assist the South African Government and selected institutions in driving the e-Research paradigm forward at an accelerated pace. It can further assist research managers in implementing an e-Research strategy within their own organisations. Lastly, the framework could assist in the development of strategy specific to e-Research.

As part of the process of developing the framework, it was possible to come to some conclusions that are perhaps not directly part of the framework. These are discussed in some detail in the next section. Similarly, the recommendations were identified and reflected in section 6.3.

6.2 Conclusions

In setting out to determine a framework for the implementation of e-Research in South Africa, four objectives were selected. These objectives related to key themes identified within the research statement, namely identifying those components that should be included in the framework, and how these various components are funded, in order to elaborate and extend on the existing framework and further calibrating this framework with at least one research institution. The findings from each of these objectives presented in Chapter 5 provide a good foundation to support the research/problem statement.
From the research, it can be concluded that a diversity of components make up the e-Research framework (see section 5.4.1 and 5.5). None of these should be ignored or underestimated. Implementation of e-Research on a national and organisational level could contribute towards building South Africa’s knowledge economy. It could also further assist in ensuring that South Africa remains globally competitive in research.

It was seen from the research that South Africa has largely been focused on the development of “cyberinfrastructure” (see sections 3.1 and 5.4.1.1); however, the research has highlighted a variety of other components that also form part of e-Research. It was evidenced that while researchers can benefit substantially from infrastructure, without the necessary tools and applications to access the infrastructure, the full potential of this infrastructure is possibly not being utilised or realised. To further the e-Research capabilities in South Africa, focus will need to be given to the development of products, services, and applications. The development of e-Research capabilities should be looked at holistically to ensure that all aspects are being addressed, as opposed to only focusing on certain areas.

For this reason, it was concluded that a body/committee of people would be required to both drive and to oversee the development (see section 5.4.2.5), implementation and planning of e-Research on a national level. A similar committee can be created within individual institutions to oversee the institution’s e-Research strategy implementation. This further highlights the need for a suitable organisational structure, in order to implement the e-Research Strategy as highlighted in section 2.2.1.

Furthermore, the researcher concluded that the framework (see section 5.5) can be used both at a national level as well as at an institutional level. Where at a national level, the development of a large high performance computing centre may be required, for
example, at an institutional level, access to the resources, as opposed to the development of the resource, may be required. The framework should thus be adapted to suit each institution’s needs. As indicated by one of the interviewees, the next step would be to take the framework to universities and research institutions, so that they can perhaps expand on the framework, but more importantly, so that they could plan how they would implement it.

The framework should also constantly be revised and adapted, as and when emerging trends and needs arise. An example of this was highlighted in section 5.4.2.2, where mobile connectivity and the need for a Virtual Research Environment came to light as being a relatively new need in South Africa. Mobile connectivity and cloud computing also did not form part of the initial UK and Australian models. A further example that has come to light is that of General Purpose Computation on Graphics Processing Units (GPGPU, 2011), which may form part of the infrastructure layer of the framework, shown in Figure 26. This could not be included in the study as it came to light at the end of the study, but should possibly be studied further and incorporated into the framework if necessary. In a dynamic environment, such as e-Research, technologies are constantly changing, and as such the framework will require continuous updating to ensure it remains relevant.

The framework, indicated in section 5.5, is very complex and involves many stakeholders from a diversity of institutions, and calls for extensive change on the part of the stakeholders. For this reason, a high level of change management, as highlighted in section 2.1, will be required, to bring about the necessary changes. The model will require a high level of discussions on the parts of the affected stakeholders and will require the “buy-in” from all affected institutions. Furthermore, it was evidenced from discussions that the researchers themselves may be reluctant to change the ways in which they are working. This could further inhibit the implementation of e-Research, but more
importantly, it could affect the potential value to be gained by implementing e-Research. The researcher observed that for such a large scale implementation, collaboration would be required between stakeholders on all fronts. This is inclusive of collaborative funding, collaborative development, and collaboration in conducting research.

The researcher further observed from the research that skills development and training would be essential to the implementation of e-Research. Not only would researchers need to be skilled, but also support staff, developers, managers, and IT personnel. A wide variety of skills would be required to implement e-Research, as well as the necessary infrastructure on which to conduct training (See section 5.4.2.3).

From the research, it was clear that intangible aspects, such as oversight, leadership and collaboration, as indicated in sections 2.2.2, 3.6 and 5.4.2.5, should also be included in a framework for e-Research. It was seen from the research (see sections 3.5.5 and 3.6.1) that these aspects can often contribute to the success of various aspects of the e-Research implementation process, both on a national and institutional level.

From discussions during the focus group, as well as in the ways that interviewees responded, the researcher observed that there may be a lot of “politics” involved in funding research. This appears to be due to the fact that funding is scarce and therefore competition for funding research infrastructure and resources is seen to be very high. This high level of competitiveness could inhibit collaborative development of e-Research on a national scale.

Many of the conclusions have led to recommendations on how to take e-Research further in South Africa. The following section highlights these recommendations relating to the implementation of e-Research.
6.3 Recommendations

The following section provides recommendations for implementing e-Research nationally and within institutions/organisations. These recommendations are based on the best practices as observed in the literature, as well as deductions made during the analysis of the data.

Overseas experience, particularly in the UK, and also indicated by one of the interviewees suggests that success in engaging the research community in e-Research is associated with the appointment of a ‘champion’, a well-respected member of the research community, recognised by government, the research sector and business as the visible face of e-Research (Sargent, 2006). It is recommended that an e-Research leader or leaders be appointed, to facilitate the e-Research reforms and to champion the e-Research vision. It is further recommended that an e-Research Strategic Oversight Committee be appointed, to steer the implementation of the e-Research Strategic Framework, in consultation with the research sector and other stakeholders.

Furthermore, to raise awareness and avoid expensive duplication of resources and further to avoid developing systems that cannot communicate easily with each other, it is recommended that the development of the various components making up the e-Research paradigm be done coordinated and on a collaborative basis, under the guidance of the proposed co-ordination of activities team and overseen by the strategic coordination committee. As indicated by AERIC (2006), in section 3.5.4, the benefits of collaboration are many, and include cost sharing across investments in e-Research, the potential to develop new research methods, new cross-discipline approaches, and new relationships and enhanced research.
The development of the components should also, however, be addressed at a national level. Various institutions could drive the development process but to ensure that all institutions have access to these tools, applications, products and services, any development undertaken by an institution addressing e-Research needs should serve the broader research community collectively.

One of the issues highlighted in the focus group discussions was access to information and the quality of electronic information. In line with Sargent’s (2006) recommendations, it is recommended that a data management plan be devised, prior to the data being collected. This data and information management plan should include information on access to the data, ownership of the data, what data might be preserved or discarded, and the conditions under which the data was collected or generated. This data and information management plan should then be incorporated into the data and information middleware and services functionality.

In order to realise the full potential of e-Research, researchers will be required to adapt their ways of working. This possibly calls for a culture change (see sections 2.2.2 and 5.4.2.3). It is recommended that broad-based discussions be held with researchers, managers and all stakeholders indicating the benefits of e-Research, and that a full-scale culture change initiative be started, on both a national and institutional level. In line with this culture change, it is recommended that a skills development and training initiative be established, to address all skills and training requirements.

Funding may be one of the biggest barriers to overcome when wanting to implement e-Research, but more importantly, it was evidenced that the mind-set of managers, researchers and people involved in implementing e-Research may need to change to be more focused on collaboration and looking at the holistic high level picture of e-Research, and possibly less focused only on how they individually are
affected, in order to bring about the necessary change. This change needs to be driven and incentivised by the management of the various institutions and through funding by government.

With regard to funding and in line with Appelbe’s (2010) findings and the research results it is recommended that partnerships be established to fund e-Research activities. These partnerships should be comprised of university/research institutions, government, and industry funding. This calls for the review of funding policies and Australia’s example of funding policies, where development must be done on a collaborative, national and non-exclusive basis, could work for South Africa as well. Furthermore, when funded with public money, all development should serve the research and innovation system broadly, and not just the host/funded institutions. It would therefore be essential to review funding policies. These should be required to encourage collaboration and co-investment. As it was seen that a clear divide between government and institutional funding cannot be drawn, it is recommended that mechanisms that encourage funding collaboration are established as an essential first step.

From the set of recommendations provided above, it is evident that further research would be required. Recommendations for further studies could include:

- A study possibly aimed at developing a collaborative funding model for e-Research, possibly starting from a policy development perspective;
- A study into the policy changes required to implement e-Research nationally;
- Each component identified in the framework possibly requires further study to unravel the complexity and requirements for implementing the component with or without collaborative funding;
- Mobile Access to products, services and infrastructure, and how these have been incorporated in the UK and Australian e-Research activities.
• A study into the role of General Purpose Computation on Graphics Processing Units in e-Research.

From the recommendations above for further studies, it can be seen that e-Research is a new and exciting paradigm, with many possibilities for researchers to study further.

6.4 Closing Remarks

This dissertation presents a framework for the implementation of e-Research in South African organisations. It addresses all the components which form part of the e-Research paradigm, and touches on the funding of the various components and how funding can affect the implementation of e-Research. Furthermore, recommendations for implementing e-Research are indicated for managers and all stakeholders to consider when implementing e-Research, both nationally and within their respective organisations. From the above, it can be seen that e-Research can pave the path forward for researchers in South Africa.

Considering all the dynamics involved in implementing the e-Research framework, it can be seen that it will be a grand challenge, but also that the benefits for South Africa are limitless. Implementation of the e-Research framework aims to revolutionise the way in which research is conducted, and to drive South Africa forward in a globally competitive market.
REFERENCES


Adams, D. Chief Director: Emerging Research Areas & Infrastructure, Department of Science & Technology. 2011. Personal Interview. 06 April, Pretoria.


[Accessed 13 December 2010]


Available from: http://sites.google.com/site/greencomputingproject/event/conference/issr-10
[Accessed 07 January 2011]

Available from: <http://www.ja.net/>
[Accessed 16 December 2010]

JISC. 2011. ‘Projects, Programmes & Services’ [online]. Bristol: HEFCE.
Available from: <http://www.jisc.ac.uk/whatwedo.aspx>
[Accessed 12 December 2010]

[Accessed 15 December 2010]


Lord, P. and Macdonald, A. 2003. Data Curation for e-Science in the UK: An Audit to establish requirements for future curation and preservation. Prepared for the JISC Committee for the support of research.
[Accessed 22 January 2011]


Page-Shipp, R.J. Retired SARIS project team leader and part time lecturer in Information and Knowledge Management at GIBS. 2011. Personal Interview. 28 March, Pretoria.


APPENDICES

Appendix 1:
Focus Group Discussion Template

AGENDA

1. Introduction
   • Introduction of researcher
   • Aim and Format of the Focus Group
   • Conventions (Confidentiality & Recording of Session)
   • Personal Introduction by participants

2. Discussion Topics
   Main components making up the e-Research Paradigm
   Components implemented in South Africa
   Discussion of framework
   Funding of Components of e-Research

3. Summing Up
   Thank you
OVERVIEW

The Framework (shown in Figure 27) is built up of layers. At the bottom, there is the infrastructure layer (coloured in blue) on top of which all other various layers are built. This infrastructure layer consists of all the physical hardware necessary to implement e-Research. On top of the infrastructure layer is the middleware layer (coloured in purple). The middleware layer comprises components necessary to allow applications to interact with the underlying hardware in the infrastructure layer. Building on top of the middleware layer, is the applications layer. The applications layer consists of various tools, applications and access to various information sources.

The three layers described above can then all be accessed by the users through a human interfaces layer. Through all layers skills development is required to ensure all researchers, librarians and IT specialists can support and participate in e-Research.

It is further indicated that two bodies should possibly be established. The first body would be responsible for the co-ordination of all activities through the various layers of the framework. It would ensure the cost-effective development, implementation and integration of the various layers. The second body would be responsible for oversight, leadership and providing strategic direction. These two bodies would be funded through the Department of Science and Technology, and would be responsible for establishing the various projects to delivery each of the components of e-Research.
Figure 27: Preliminary e-Research Implementation Framework
DESCRIPTION OF COMPONENTS

1. Oversight & Leadership Committee
A committee to provide the strategic direction, drive and engagement and coordination of effort of research groups involved in e-Research. Leadership refers to the appointment of an e-Research ‘champion’ or leader, a well-respected member of the research community, recognised by government, the research sector and business as the visible face of e-Research (Sargent, 2006).

2. Co-Ordination of Activities Management Team
A team to ensure co-ordination of activities, across the various layers of the framework. To be cost effective nationally, all activities should be co-ordinated in their planning, development and implementation.

3. Grid/Cloud Infrastructure Layer
Grid is described by Foster and Kesselman (1999) as being a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities. South Africa has also invested in building Grid infrastructure and has established the South African National Grid. The South African National Grid is a federation of resources, institutes and Virtual Organisations. It is a platform for collaboration and research, providing many services to scientists (South African National Grid, 2010). According to the South African National Grid (2010), the national grid forms part of an e-Infrastructure and provide support to users of scientific computing and collaboration.

• National Backbone Network: A network for all provincial and inter-institutional networks to connect to, as well as connecting internationally to other education and research networks and grids, such as in Europe, Australia and America. The Research and education network in South African is called SANRen (SANRen, 2011).
• **Provincial Networks**: A network linking all institutions in a province and connecting to the national backbone network.

• **Inter-Institutional Networks**: A network linking all institutions together, such as all universities in a province, and linking research institutions. Then linking these separate networks to a provincial network.

• **High Performance Computing**: Researchers now need access to powerful high performance computing capacity (NRIC, 2006). They require services such as grid computing and access to databases that make increasingly more extensive use of high performance computing facilities in a collaborative environment (NRIC, 2006). South Africa has developed the Centre for High Performance Computing (CHPC, 2010) in Cape Town. Through using this system, the CHPC “aims to enhance significant research, address grand challenges, and grow computational research” (CHPC, 2010).

• **Digital Repositories**: This is where digital content, assets, are stored and can be searched and retrieved for later use. A repository supports mechanisms to import, export, identify, store and retrieve digital assets (JISC, 2011). In South Africa, the “Very Large Database (VLDB)” is in the development phase (CHPC, 2010). It is aimed at complementing the Centre for High Performance Computing through the effective curation of large databases, specifically in areas of environmental and climate change modelling and astronomy (CHPC, 2010).

4. **Middleware Layer**
   A major development of grid computing has been the development of middleware, the communications layer that allows applications to interact across hardware and network environments (Paterson, Lindsay, Monotti & Chin. 2007).

   • **Grid Middleware & Services**: Provides essential access, communication, accounting, security, trust, and co-ordination
services between the (computational and data) resources of the grid and the higher-level services that use them (DSTC, 2004).

- **Data and Information Middleware & Services**: Provides tools and services that enable the indexing, archival, discovery, analysis, integration, management and preservation of large heterogeneous distributed data repositories and digital archives (DSTC, 2004).

- **Knowledge Management Middleware & Services**: The focus of this middleware is to make use of data and information generated, archived and indexed to mobilise knowledge and aiming towards building a knowledge economy.

- **Collaboration Middleware & Services**: Provides tools and services to support formal and informal, real-time and offline collaborative activities between remotely located researchers, research communities, and resources (dynamic, scalable virtual organizations) (DSTC, 2004).

5. Access to Information

Scientific and scholarly publishing is now evolving along two distinct paths – one in which large commercial publishers are increasing their dominance in such areas as ‘branded’ journal titles and access to scientific publication, and the other in which there are a variety of open access initiatives (Houghton, *et al.*, 2003).

- **Commercial Publishing**: The pursuit of high quality research requires ready access to the published work and data of other researchers, and the facility to publish one’s own findings (Page-Shipp, *et al.*, 2005).

- **Open Access Data**: In the traditional commercial model, payment of a licence or other access fee is required, whereas an emerging model is of ‘free’ or ‘open’ access to online publications in formal journals or institutional or national repositories (Page-Shipp, *et al.*, 2005).
• **Primary Data Sharing:** Making primary research data available to other researchers is part of the e-Science paradigm. One needs to store not only the database *per se*, but also sufficient metadata to enable a potential user to find relevant data and be satisfied with its value and provenance (Page-Shipp, *et al.*, 2005).

6. **Tools, Applications & Services**

• **Visualisation:** Visualisation can be divided into two broad types or categories; Scientific Visualization and Information Visualization (Charters, 2010). According to Charters (2010), scientific visualization is concerned with the conversion of numeric data, usually from scientific experiments and simulations to a graphical representation. Information visualization on the other hand is concerned with the conversion of other forms of data, structured and unstructured text, images and video, to an appropriate graphical representation.

• **Access, Authentication & Authorisation:** Security technologies must support a variety of user and provider requirements such as authentication, authorisation, trust, privacy, policy management, and information assurance, all in a user- and provider- friendly framework (ISSR, 2010).

• **Communication & Collaboration:** Collaboration may involve remote control of equipment, sharing of data and access to content in repositories and increasingly communication and conferencing (Sargent, 2006), including audio conferencing, video conferencing and instant messaging.

• **Digital Curation & Preservation:** This activity refers to the active management of datasets for their scientific and scholarly useful lifetimes, including the promotion of effective and widespread use. Other ePrint archives, for example, theses and research reports, form part of this category (Page-Shipp, *et al.*, 2005).
• **E-Learning/Digital Scholarship Tools & Applications:** Tools and Applications to provide an innovative learning experience for learners (Li, *et al.*, 2006)

• **Project Specific Tools & Applications:** For specific research projects, certain tools may be required and interfacing with these tools will be necessary, for example, telescopes.

7. **Applications/Human Interfaces**
Applications to make the underlying modules and data accessible and easy to use and access.

8. **Skills Development**
A deeper and more rapid up-take of e-Research would require a considerable increase in the number of researchers with the confidence, knowledge and technical skills to undertake e-Research, and in the number of IT specialists who can support or participate in e-Research projects with domain-based researchers (Australian Government, 2006).
Appendix 2:

An Implementation Framework for e-Research: Interview Schedule

Introduction by researcher:
This interview forms part of a current research study being conducted for the fulfillment of requirements for a Masters Degree in Business Administration through Unisa’s School of Business Leadership. The researcher, Shelly Fernihough, is investigating the components making up the e-Research paradigm and how these components are funded. The Interview is designed to obtain feedback from you based on your experience and knowledge of e-Research.

Would you have any objections to this session being recorded?

For the purpose of referencing, would you grant permission for your name and the details of this session to be recorded, to be used as references?

Interview Questions:

1. From the framework, are there any components that make up the e-Research paradigm that have not been included?
2. From the framework, are there any components that have been included but do not form part of e-Research?
3. In your opinion what are the weaknesses and strengths of the model?
4. How should e-Research be funded and by who? Centralised or de-centralised.
5. Which components should be funded nationally by government and which by individual institutions?
6. Looking at the revised framework, what are your opinions of the changes made?