Mousaion

Contents

ii

DETERMINING THE REASONS WHY ICT CENTRES FAIL: SIX SOUTH AFRICAN CASE STUDIES

DAVE HULBERT

Department of Information, Communication and Technology, University of South Africa, PO Box 392, Pretoria 0003, South Africa
DHulbert@tsa.ac.za

MARITHA SNYMAN

Department of Information Science, University of Pretoria, Pretoria 0001, South Africa
maritha.snyman@up.ac.za

ABSTRACT

It is generally believed that ICT centres will provide developing communities with access to digital information that will place them be in a position to bridge the digital divide. Yet there is little evidence of the successful establishment of ICT centres in South Africa. There appears also to be no evidence of a methodology or structured approach for measuring the effectiveness of ICT establishment. This article discusses the application of a proposed approach and measuring instrument that can be used for assessing the effectiveness of the ICT centre establishment. The outcomes of this first longitudinal study of six ITC centres that were established in South Africa highlight the reasons why ICT centres fail and suggest that there are deeper reasons for this failure.

KEYWORDS

ICT centres, telecentre, action research, methodology for, South Africa, failure of, longitudinal study, measuring instrument

1   INTRODUCTION

Telecentres, (called ICT centres for the purpose of this article) are universally seen

1

as the solution to universal access simply because full, universal service (a phone in every home) is not a realistic goal for developing countries (ITU 1998). Using ICT centres for providing access to telecommunications and information in developing countries and/or communities is also referred to as the "centre approach" (Snyman & Snyman 2003).

ICT centres have also been described as "... powerful engine(s) of rural development" and "(a) preferred instrument(s) in the fight against poverty" (World Bank 1998). The view that access to ICTs can solve many social and developmental problems is one of the reasons why the centre approach is heralded as the solution to the ever-widening digital divide. Ernberg (1998) describes, for instance, how rural community centres can empower communities to become part of the information society and argues that access to ICTs could be vital to the economic development of disadvantaged communities. African and Asian case studies that prove these points are described by, amongst others, Kiplang'at (1999).

Faced by a skewed information communication infrastructure after the ANC came to power in 1994, the new government expressed the belief that ICTs can be a positive force for social change, that " there is a direct, positive correlation between access to telecommunications and socio-economic development …" (Naidoo 1997) and that "the modern communication technology ... must help us educate our children, particularly in the rural and other underdeveloped areas of our country" (Mbeki 1998).

Problems such as poverty, and a poor electricity and telecommunications infrastructure, as well as low education levels and a lack of computer skills, stood in the way of reaching these ideals. Faced with such problems, the South African government subscribed to the centre approach as a strategy to place ICT and other information services within reach of the disadvantaged rural South African communities. Various initiatives[1] were undertaken to establish ICT centres in the developing communities of South Africa. These include the telecenters of the Universal Service Agency (USA), the Multi Purpose Community Centres (MPCCs) of the Government Communication and Information Services (GCIS) and the Community Post Offices (CPOs) and Public Information Terminals (PiTs) of the South African Post Office. Not only did the government embrace this approach --- various donor agencies, private enterprises and other state and parastatal institutions also embarked on the implementation of ICT centres in developing communities.

 Unfortunately, however, ICT centres in South Africa have not been very effective in securing space for its developing communities on the information super highway. The

1   Although these initiatives differ in size and structure they can all be regarded as ICT centres because they aim to provide community access to ICTs (See Benjamin et al 2000 and Snyman & Snyman 2003 for a typology.

inability of ICT centres to effectively link developing countries, in general, (Dasgupta & Gopalakrishnan 2002; Heeks 2002) and developing communities in South Africa, specifically, (Conradie 1998; Van Audenhove 1999, Snyman & Snyman 2003) to ICTs has been described in various sources. Many ICT centres established in developing countries have been forced to close down owing to their ineffectiveness and the lack of sustainability (Michel 1997; Sassen 2002). In South Africa, telecenters of the Universal Service Agency have been struggling to survive and, after years of inefficiency, have been taken over by Sangonet in an effort to revive the initiative (Stones 2007). The success rate of MPCCs regarding technology and information transfer is also poor (Van Audenhove 1999; Stilwell et al 1999; Benjamin et al 2000; Snyman & Snyman 2003; Holmner & Snyman 2006). Those ICT centres that have not closed down often have to rely on handouts from donors for their continued existence (Mphahlele & Maepa 2003*).*

2 Background

In the apparent absence of a methodology or structured approach for measuring the effectiveness of ICT establishment, this article proposes a research approach and measuring instrument that can be used to assess the effectiveness of ICT centre establishment. The research approach and measuring instrument are applied in a study of six ITC centres. The findings of this study highlight the reasons for the failure of the six ICT centres.

Most South African studies investigating the effectiveness of ICT centres evaluate the status of selected ICT centre at a specific point in time (Benjamin *et al* 2000; Mphahlele & Maepa 2003; Snyman & Snyman 2003). To date, there has been no attempt to trace the life cycle of ICT centres over a period of time. In a longitudinal study it is possible to determine how the implementation and management of ICT centres contribute to their failure and/or success. To be able to evaluate the same ICT centre(s) after a lapse of time, a standardised evaluation instrument had to be developed.

To achieve its aim, the study focused on attempts of the former TechnikonSA (TSA) and various partners to implement six ICT centres in certain rural developing regions of South Africa. In this first longitudinal study, the establishment and progress of six ICT centres were traced over a period of four years - from 1998 to the beginning of 2006. The TSA supplied the capital and human resources for the establishment of the ICT centres and, in some instances, managed additional funds (ie donations). The TSA was responsible for the implementation of the ICT centres and the donors relied on the TSA to effectively manage their donations for the implementation of the ICT centres.

3

3 An action research approach

As the study was to span four years, it required a long-term interaction between the researcher and role players involved in the ICT centres. An action research approach was therefore chosen.

Action research was developed during the 1960s and is particularly useful in the area of managing change (Alan 1991). Action research is seen as a long-term process that involves both intervention and research (Reason & Bradbury 2001). It usually involves a small-scale intervention on the part of the researcher in the phenomenon that is being studied. The action researcher becomes actively involved while doing the research (Alan 1991). French & Bell (1990) define action research as "the process of systematically collecting research data about an ongoing system relative to some objective, goal or need of that system; feeding these data back into the system; taking action by altering selected variables within the system based both on the data and on hypotheses; and evaluating the results of the actions by collecting more data".

Effective action research presupposes quality access to the research situation. This, in turn, makes for a demanding data-collection process. Since the researcher often has to fulfil the dual role of consultant and academic researcher (Bell 1993), objectivity issues come into play. Action researchers doing research in developing countries are, for example, often criticised by social scientists for three reasons (Bless & Higson-Smith 1995):

• Extraneous project variables are difficult to control; results are therefore sometimes based on the views of the researcher.
• Objectivity can become an issue owing to the close relationship between the researcher and research subjects.
• Community-specific problems may narrow the focus of the research, which prevents generalisation of the findings to other communities.

For this study to succeed, it was necessary to gain the cooperation of the personnel of TSA and the ICT centres, to fulfil the dual role of consultant and academic researcher and to design a research process that would, as far as possible, eliminate subjectivity. These challenges were addressed by employing various triangulation measures, which included three data-collection methods and six comparable case studies. A standardised evaluation instrument was developed to determine the progress of the ICT centres according to predetermined criteria in an effort to increase objectivity.

4 Data-collection methods

Henning (2004) mentions three data-collection methods used in action research: par-

4

ticipant observation, interviews and document analysis. All three of these methods were used.

Participant observation

The aim of participant observation here was to gain an insider's understanding of what is being researched (Denzin & Lincoln 2002). The researcher "immersed (himself) in the day-to-day activities of the community being studied … in order to provide a record of observed behaviour under varying conditions" (Meyer 2000) and aimed to objectively record the data without influencing the responses of the research popula-tion. Participant observation was specifically used to try to assess the effectiveness of the ICT centres and the commitment of the role players involved in the management of these centres.

Interviews

Semi-structured interviews were used, since this enabled the interviewer to alter the questions during the interview itself (Bless & Higson-Smith 1995; Merriam 1988) while attempting to attain a measure of objectivity. A list of questions was used as a guideline, but the exact order of the questions was not pre-determined to allow the interviewer more freedom in the data-collection process.

The limitations of this process lie in the possibility of dishonest responses and the possibility of a biased interpretation of these responses (owing to the relationships which inevitably develop between interviewees and researcher over a period of time). The researcher attempted to counteract this problem by involving third parties in the interviewing and analysis process.

Document analysis

Documents used for analysis in this study included both primary and secondary (Finnegan 1996) documents. All available documentation was sourced from the differ-ent ICT centres. This included contracts, correspondence with stakeholders, registers, motivations and project plans, invoices and business plans.

The limitation here is that the interpretation of the documents can be influenced by the researcher's aims and point of view. This was avoided as far as possible by using triangulation in the data collection and analysis.

5. A case study design

Each of the ICT centres included in the study was regarded as a separate case study. The

5

case study methodology is used as an umbrella term for a family of research methods which have in common the decision to focus on an inquiry around a specific instance or event (Bell 1993). A case study can therefore be defined as an empirical inquiry that investigates a phenomenon within its real life context and in which multiple sources of evidence are used. The following four applications of a case study design listed by Yin (1994) were all relevant for this study:

- To explain complex causal links in real-life interventions.
- To describe the real-life context in which the intervention has occurred.
- To describe the intervention itself.
- To explore those situations in which the intervention being evaluated has no clear set of outcomes.

Every one of the six ICT centres in this study was considered a separate case study and was subjected to the same investigation process. The results of the separate case studies were compared and integrated to provide a comprehensive picture of the implementation process of the ICT centres.

The following steps were followed in the investigation of each ICT centre:

- Determine the status quo of each ICT centre.
- Formulate hypotheses based on the first step.
- Suggest adjustments based on the different hypotheses.
- Re-evaluate after a period of time.
- Draw conclusions based on the findings of the re-evaluation.
- Compare and integrate findings.

An evaluation instrument was developed based on previous research and a literature review. The instrument was systematically applied to each centre during steps 1 and 4 of the investigation process. This systematic application of a standardised instrument helped to counteract the problem of subjectivity. The measuring instrument is described in more detail in the next section.

6 The evaluation instrument

An evaluation instrument was designed that was flexible enough to cater for the peculiarities of each centre, but also able to evaluate the criteria that were common to all the ICT centres.

The evaluation instrument consisted of the following four broad phases:

1 Conceptualisation
2 Implementation
3 Evaluation and adjustment
4 Sustainability

6

The conceptualisation phase examined the motivation for the centre and the reasons for deciding on the physical location. It also examined the community needs that were to be addressed by the ICT centre. The implementation phase dealt chiefly with the actual establishment of the ICT centre and the selection of the role players responsible for the centre's functioning. This phase also examined the impact of the involvement of third parties in the implementation of the ICT centre, including the project manager's role. The evaluation and adjustment phase concentrated on the effectiveness of the ICT centre. The suggested changes and the associated evaluations made by the researcher were recorded in this phase. The sustainability phase established the potential of the ICT centre's ability to operate independently without additional resources. The sustainability phase depended on the continual evaluation of the previous phase.

Each one of the four phases was broken down into smaller units that were then used to determine the metrics for the evaluation process. Each of these phases is now discussed in more detail.

The conceptualisation phase

The conceptualisation phase covered the process that was followed in planning the implementation of each ICT centre. Table 1 below provides a "bird-eye's view" of the units, their associated evaluation and the data sources used.

UNIT[1]	MEASUREMENT[2]	DATA COLLECTION METHOD[3]
Site selection	Needs of the community	Documentation
	Size of community	Documentation
	Access to other ICT centres	Interviews with ICT centre staff; document analysis
	Existing infrastructure	Interviews with centre staff; document analysis
	Access to centre	Interviews with centre staff and community members & document analysis
Role players	Community participation	Interviews with centre staff and community members
	Profile of role players	Interviews with TSA role players and local leadership
	Acceptance by community	Interviews with centre staff and community members
	Profile of centre management	Interviews with ICT centre staff and TSA role players

Community needs	IT training	Interviews with role players[4]
	Telecommunications	Interviews with role players
	Internet access	Interviews with role players
	E-mail	Interviews with role players
	Information needs of community	Interviews with role players
Business plan	Funding: Proof of seeding money	Document analysis
	Income generation: Forecast criteria used	Document analysis
	Marketing: evidence of plan	Document analysis
	Support and maintenance: Proof of contracts/service level agreements	Document analysis

Table 1: Conceptualisation phase of the evaluation instrument

Site selection

Site selection refers to the reasons for deciding upon a physical location for an ICT centre. In order to understand why the physical site was chosen, the following criteria were used:

- The needs of the community that could be addressed by an ICT centre and the documentation prepared by the TSA to justify this choice of site were used as resources.
- Data regarding the size of the community that would benefit from the ICT centre was obtained from TSA's project proposal and other documents that contain information about the demographics of the area.
- Alternative points of access to ICTs in the area were determined through interviews with the personnel of each ICT centre.
- Data regarding the existing infrastructure, such as access to telephones, power, transport and the physical building housing the centre was obtained through document analysis and interviews with the personnel of each ICT centre.
- Information about the physical accessibility of the centre for community members was obtained through interviews with community members and ICT centre personnel.

Role players

Role players include all people who had a hand in the implementation and management of the centre, as well as those who would benefit from the centre. Community participation refers to the extent that the community was involved in the initial stages of the centre implementation. This was determined through interviews with the centre's staff and community members.

A profile of the role players in the project was drawn up using data collected through

8

documentation and interviews with the project manager and ICT centre personnel.

In an attempt to establish whether there was community ownership, the acceptance of the centre by the community was determined by interviews with the centre staff and community members themselves. The profile of centre management was determined through interviews with the project manager and other decision makers.

Community needs

The intention was to establish which community needs could be addressed by the services offered by the ICT centre. Services usually include access to the Internet, IT training, telecommunications, copying and typing. This data was obtained through interviews with community members and those responsible for identifying the location of the centre and the documentation motivating the establishment of the centre.

Business plan

The business plan was used to assess how those responsible for the conceptualisa-tion of the centre had planned for the financial viability of each ICT centre. This was determined by analysing the project proposals, business plans (if available) and other related documentation. The following points were noted:
- Funding and how the funding was sourced.
- Income generation and how this was to be achieved over the medium to long term.
- The marketing plan.
- Support and maintenance of equipment used in the centre. Documents such as agreements with third parties served as additional sources.

Implementation phase

A description of the implementation phase, the next step in the measuring instrument, is preceded by a summarised presentation in Table 2 below.

UNIT	MEASUREMENT	DATA COLLECTION METHOD
Selection of centre staff	Business skills	Interviews with staff of TSA and ICT centres
	Technical skills	Interviews with staff of TSA and ICT centres
	Knowledge of community	Interviews with staff of TSA and ICT centres
	Acceptance in community	Interviews with staff of TSA and ICT centres
	Innovation	Interviews with staff of TSA and ICT centres
Implementation process	Project plan	Documentation; interviews with the project manager

9

IT suppliers and IT support	Process followed	Documentation analysis; interview with staff of ICT centre staff and role players of TSA
Involvement of community leaders	Process used	Documentation analysis; interview with staff of ICT centre staff and role players of TSA
Identification of other support structures	Process used	Documentation analysis; interview with staff of ICT centre and role players of TSA

Table 2: Implementation phase of the evaluation instrument

Selection of centre staff

This refers specifically to the selection process used for the appointment of staff responsible for managing the ICT centres. Criteria considered important in the selection process included business skills, technical skills, local knowledge, acceptance by the community and the ability to think laterally or innovatively. This data was collected through a series of interviews with those responsible for the actual implementation of each centre.

Implementation process

The manner in which the centre was deployed was determined by examining the project plan or related documents. Interviews with the project manager were used to collect additional data.

IT suppliers and IT support

This refers to the selection process applied to select IT suppliers and IT support. Issues such as the process used to select suppliers, remote support, on-site support, other interests of the supplier in the area, third parties used by the supplier and the financial position of the supplier were investigated in interviews with staff of the ICT centre and role players of TSA. Documentation was also perused to establish how the IT suppliers and support were selected.

Involvement of community leaders

The identification of community leaders was examined through interviews with ICT centre personnel and the TSA. Issues that were investigated included the following: how community leaders were identified, the role of community leaders, type of agreement with community leaders regarding feedback and community leaders' expectations of the ICT centre. Documentation available from the project manager was also analysed.

Other support structures

Other support structures refer to third parties and services that the ICT centre needed to be sustainable. The existence of back-up plans for the supply of electricity and the

type of transport available for reaching the ICT centre were investigated. The data sources were the same as those used in the previous units.

7 Evaluation and adjustment phase

Evaluations and adjustments were applied on an iterative basis to each ICT centre. The effectiveness of this part of the instrument depended on the ability of the ICT centre's management to effect the changes suggested to them. The first evaluation was an attempt to ascertain the baseline situation of each ICT centre. This was used as an initial benchmark for further assessment. After the initial evaluation, modifications to processes and the management of the ICT centres were suggested if deemed necessary. The implementation of these suggestions was assessed after a reasonable amount of time in order to determine their effectiveness.

The evaluation and adjustment phase focused on four major areas: centre usage, centre reliability, the perceptions of users about the standard of the services, and income generated versus expenditure. Table 3 below reflects the third phase of the instrument.

UNIT	MEASUREMENT	DATA COLLECTION METHOD
Usage of centre	Number of people	Document analysis and interviews with staff of ICT centre
	Number of training sessions	Document analysis and interviews with staff of ICT centre
	Internet usage	Document analysis and interviews with staff of ICT centre
	E-mail usage	Document analysis and interviews with staff of ICT centre
	Other IT services	Document analysis and interviews with staff of ICT centre
Centre reliability	Up time of server	Document analysis and interviews with staff of ICT centre
	Up time of workstations	Document analysis and interviews with staff of ICT centre
	Connectivity to the Internet	Document analysis and interviews with staff of ICT centre
	Power availability	Document analysis and interviews with staff of ICT centre
	Access to third party suppliers	Interviews with centre staff

11

Assessment of user perceptions	Affordability	Registers, accounts, interviews with staff of ICT centre and community members
	Appropriate services	Document analysis and interviews staff of ICT centre and community members
	Ease of centre accessibility	Document analysis and interviews with staff of ICT centre and community members
	Operating hours	Document analysis and interviews with staff of ICT centre and community members
	Awareness of centre	Document analysis and interviews with staff of ICT centre and community members
Income generation	Amount	Document analysis and interviews with staff of ICT centre
Overheads	Amount	Document analysis and interviews with staff of ICT centre

Table 3: Evaluation and adjustment phase of the evaluation instrument

Centre usage

This unit relied on documentation that indicated the number of people using the centre, the number of training sessions offered and the usage of services, such as the Internet, e-mail and other services available at a specific ICT centre. Registers and interviews with ICT management staff were used as a basis for data collection.

Centre reliability

This examined the actual availability of services provided by each centre. Reliability was also considered a function of the support that each centre received from third party suppliers and, for this reason, access to third party suppliers was considered important. Server uptime and workstations were measured through registers kept by the centre and interviews with centre staff. Access to the Internet, telephones and power, as well as back-up measures (eg the presence of generators and UPSes) were also mainly determined through interviews with ICT centre staff and available documentation pertaining to the status of access to these services.

User perception of the ICT centre

All the measures, such as affordability, appropriate services, ease of access to the centre, operating hours of the centre and awareness of the centre were determined through interviews with ICT staff and community members. Often, particularly in the rural areas,

12

for a variety of reasons it was not possible to interview community members and the researcher had to rely on feedback from the centre staff about user perceptions.

Income generation and expenditure

This was determined through the documentation made available by the staff of the ICT centre in which income generation and expenditure were recorded. Interviews were used when the documentation was not up to date.

Sustainability

Sustainability was determined by the collection and analysis of the data obtained during the evaluation and adjustment phase.

Sustainability of a centre was determined by measuring four factors: centre usage, income generated, expenditure and user perceptions. These factors were identified over the period of the research study and the intention was to note any changes in order to determine growth or improvement. This was done by comparing figures with previous measurements.

Analysis of the data

The data collected from each ICT centre was analysed by content analysis and organised into comparable tables based on predetermined categories by the evaluation instrument. An example of the way in which the data was coded is presented in Table 4 below.

SITE SELECTION	RESPONSE
Existing infrastructure	
Does the site have access to electricity? If not are there alternatives?	No. One petrol generator No
Does the site have access to water, if not how is water accessed?	Yes
Are there ablution facilities?	Yes
Does the site have access to telephone lines?	Yes, one line
Where is the ICT centre housed?	In a secure building
Is the site secure from theft?	Yes
Is the site safe from the elements?	Yes
Access to centre	
Is the site close to or part of a community centre?	Yes
How is it patronised by the community	Numbers are growing but is still difficult to ascertain
Is the site easily accessible by the community?	Yes. On the main bus and taxi routes

Table 4: Site selection of ICT1

13

Similar tables were drawn up for all the phases evaluated by the evaluation instrument, as well as for each of the six ICT centres. These factors were used as indicators of each ICT centre's sustainability potential. Their sustainability potential was deduced from the sum of these factors.

8 Findings

Table 5 below is a summary of the status of the six ICT centres at the end of the research process.

ICT CENTRE	EFFECTIVENESS
ICT1	In the first years ICT1 was not often used Slight growth occurred after the community centre it is part of, received an international award Currently not effective
ICT2	Slow start Showed real growth during the last two years Used by community Income generation exceeds expenses Is realizing potential, but not yet independent
ICT3	Good start Consistent negative growth No longer effective Has not realised potential
ICT4	Non-starter due to change of location just before implementation Ineffective at current location due to low patronage Has never shown any potential
ICT5	Poor start Stagnant Has never realised potential
ICT6	Slow start Has hardly been used in the past four years Has never realized potential

Table 5: Summary of the six ICT centres' status at the end of the research process

Five out of the six attempts to effectively deploy ICT centres proved not to live up to the expectations of the initiators. The only ICT centre that did show signs of becoming sustainable was still unable to operate completely independently and had to rely on some support from the TSA.

The findings relating to each ICT centre were integrated and compared with present conclusive findings of the comprehensive action research process. Factors that contributed to the failure of the ICT centres were examined and compared with findings of similar studies and literature on the topic. The results of this process are discussed below.

14

Reasons for failure

The following are the main reasons why ICT centres fail.

- No community needs analysis was undertaken for any of the ICT centres. The community needs were simply assumed by the initiators. This made it more probable that the services provided by the ICT centres would not meet the community's needs. In short, communities seemed to have been given a solution to a problem that they themselves were not aware of. In fact, one of the main reasons for the low success rate of ICT centres is that developers have "solved" problems that developing communities simply do not regard as problems (Schoen 2002; Heeks 2002; Colle & Raul 1999).

- No landscape audit was carried out to ascertain the feasibility of the selected locations in terms of electricity supply, access to telephone lines, local culture, legislation and other issues which could impact on the success of the ICT centre. When establishing ICT centres in developing areas, much has been documented regarding the importance of understanding local conditions with respect to infrastructure, legal requirements, access to communication and the Internet. There is general consensus that the obstacles, challenges and profile of a community must first be identified and analysed so that this information can be used as input for an ICT centre framework. It is also clear from the literature that conditions in seemingly similar environments may well differ significantly from each other and that each project aimed at setting up an ICT centre should be considered unique. This means that the project team responsible for the establishment of the centres must be fully aware of the prevailing conditions and adapt their respective approaches to ICT implementation accordingly. There is overwhelming evidence throughout the developing world that the rendering of appropriate services is the key to long-term viability and sustainability (Sabien 2002; Fuchs 2000; Thamizoli & Balasubramanian 2001; Conradie 1998; Snyman & Snyman 2001; Kirkman, 1999; Benjamin et al 2000).

- Those who implemented the ICT centres did not involve the communities or authorities in the planning and implementation of the ICT centres. Since the communities were not engaged in the process, there was obviously no community buy in. Communities' ownership of, and involvement in, ICT projects, are a prerequisite of the success of these projects. Communities that are simply presented with an ICT centre as a *fait accompli* often show no involvement (Conradie et al 2003). An ICT centre must be established with the full participation of the local community (Khumalo 1998). In short, the 'box drop' approach is a recipe for failure (Benjamin et al 2000; Snyman & Snyman 2003).

- Contrary to recognised good business practices, no business plans or financial forecasts based on community needs were produced during the planning or implementation phase. Nor was there any guidance given to those left to do the day-to-day managing of the ICT centres.

15

- No partnerships with service providers, local government or civil organisations were formed that could have helped with the utilisation and maintenance of the ICT centres. The three ICT centres that did attempt to form partnerships showed an improvement in service delivery at the time. The only two ICT centres that were still operating were those that had, in fact, continued to be part of a partnership. The importance of identifying and forming strong partnerships when dealing with advancement projects in developing countries was again highlighted in the final report of the outcomes of the World Summit on Sustainable Development, which was hosted in Johannesburg in 2002 (Whelan 2002; Kyle 2002). There is strong evidence to suggest that ICT centres can only survive through partnerships and being part of the community (Kyle 2002; Fuchs 2000; Sabien 2000).
- Little, if any, training of ICT staff members formed part of the planning and implementation process. Because of this lack of training, ICT centre staff had no relevant technical expertise and business skills. The literature emphasises the fact that the feasibility of ICT use in the developing world is only possible through training and education (Kirkman 1999). The staff responsible for the management of the ICT centre must continually be trained to keep abreast of developments in software, hardware and networking technology. Identifying and training local champions, who will nurture the ICT centre project, can make or break the success of such a service (Gómez & Ospina 2002; Fuchs 2000).
- The same implementation model was repeatedly used by TSA even after it has been proven to fail, and despite the fact that theory and previous research warn against the method of implementation that was followed here. In addition, four of the six ICT centres were implemented by the same individual over a number of years. Peled (2001), Clark (2002), Myers and Young (1997) and Berg (1993) all refer to this problem. They suggest that continued attempts, by organisations and individuals, to transfer ICT to developing communities through ICT centres could be the result of hidden agendas on the part of these individuals and organisations (simply because these projects have shown so little evidence of success).

It is clear that the failure of the ICT centres established by the TSA and its partners was predictable. All the reasons for failure revealed by this research have previously been explicitly addressed in the literature.

9 Conclusion

The outcomes of this first longitudinal study, using an action research approach, not only corresponds with what has been documented by previous research (about the reasons for the failure of ICT centres), but point to more complex reasons for the overall failure to implement and sustain ICT transfer in developing communities. The

16

factors for the failure of ICT centre implementation in Southern Africa indicate, in no uncertain terms, the importance of the community as the most feasible departure point. These findings are unequivocally supported by other research and by various theories relating to participatory communication, information dissemination and technology transfer. Yet the public, as well as the private sector, happily continue to use the same ineffective approaches when deploying ICT centres (Mphahlele & Maepa 2003; Snyman & Snyman 2003, Benjamin et al 2000; Stones 2007).

It is not at all clear why approaches that are known to fail continue to be used. The role of the donors and consultants and their motivation for implementing an ICT centre in a developing community if there is no ownership or any long-term interest for the donor or consultant involved have been pointed out in this study. This study revealed that the involvement of third parties was characterised by:
•   A top-down, modernistic approach to development.
•   "Throwing technology" at a problem that is a symptom of more profound problems such as poverty, unemployment, lack of infrastructure, education deficiencies, etc. --- problems that are probably more difficult to address.

The public sector's track record regarding the implementation of ICT centres in South Africa is not much better. Similar implementation and management approaches provide similar results in ICT establishment by the public sector (Benjamin et al 2000; Snyman & Snyman 2003; Holmner & Snyman 2006). MPCCs are often launched with great fanfare and then fizzle out for the very same reasons discussed in this study (Stones 2007.)

One has to ask whether, in fact, short-term political or monetary gain is favoured above the long-term development of communities (which might be the result of a more partici-patory approach). The existence of these underlying problems suggests more complex issues than the straightforward guidelines usually provided as "recipes" for ICT centre implementation. It is these more profound problems that need to be investigated.

References

Alan, G. 1991. *Handbook for research students in the social sciences*. London: Dalmer.

Bell, J. 1993. *Doing your research project.* Milton Keynes: Open University Press.

Benjamin, P. 2000. *Link Centre annual report.* Available: http://www.communitysa.org.za/T2000LitRev.htm Accessed 23 September 2003).

Benjamin, P, Stavrou, A, Burton, P & McCarthy, C. 2000. *Telecentre 2000: synthesis report.* Johannesburg: DRA-Development and LINK Centre.

Berg, E. 1993. *Rethinking technical cooperation: reforms for capacity building in Africa.* New York: United Nations Development Programme.

Bless, C & Higson-Smith, C. 1995. *Fundamentals of social research methods: an African perspective.* Kenwyn: Juta.

Clark, J. 2002. *Promoting participation in telecentres*. Available: http://ip.cals.cornell.edu/commdev/documents/jdc-clark.doc (Accessed 17 February 2006).

Colle, R & Raul, R. 1999. *Communication centres and developing nations*. Available: http://www.devmedia.org/document/Banga.htm (Accessed 12 August 2000).

Conradie, PD. 1998. Using information and communication technologies (ICTs) for development at centres in rural communities: lessons learned. *Communicare* 17(1):97-117.

Conradie, D, Morris, C & Jacobs, S. 2003. Using Information and Communication Technologies (ICTs) for deep rural development at centres in South Africa: research case studies. *Communicare* 29(1):199-217.

Dasgupta, S & Gopalakrishnan, S. 1999. Determinants of information technology adoption: an extension of existing models to firms in a developing country. *Journal of Global Information Management* 7(3):30-53.

Denzin, NK & Lincoln YS. 2000. *Introduction: the discipline and practice of qualitative research*. London: Sage.

Ernberg, J. 1998. Telecentres: access for all. *Telecommunications (International Ed.)* 32(4):10-11.

Finnegan, R. 1996. *Data collection and analysis*. London: Sage.

French, J & Bell, W. 1990. Choosing an appropriate research methodology. *Construction Management and Economics* 15(2):101-124.

Fuchs, R. 2000. Telecentres share the tools of the information age. *The Unesco Courier*, March:43 -45.

Gómez, R & Ospina A. 2002. *The Lamp without a Genie: using telecentres for development without expecting miracles*. Available: http://ip.cals.cornell.edu/commdev/documents/jdc-gomez.doc (Accessed on 16 February 2006).

Heeks, R. 2002. *Failure, success and improvisation of information systems projects in developing countries*. Available: http://www.unige.ch/iued/wsis/DEVDOT/00345.HTM (Accessed 16 October 2004).

Heeks, R. 2002. I-Development not E-Development. *Journal of International Development* 8(14):1.

Henning, E. 2004. *Finding your way in qualitative research*. Pretoria: Van Schaik. Available: http://www.iccwbo.org/home/environment_and_energy/sdcharter/topics/sustain/VE%2520report.pdf (Accessed 10 March 2004).

(ITU) International Telecommunications Union 1998. *World telecommunications development report*. Geneva: ITU.

Kiplang'at, J. 1999. An analysis of the opportunities for information technology in improving access, transfer and the use of agricultural information in the rural areas of Kenya. *Library Management* 20(2):15-127.

Kirkman, G. 1999. *It's more than just being connected*. Available: http://www.unbotswana.org.bw/undp/docs/bhdr2002/beingconnected.pdf (Accessed 11 November 2002).

Kyle, R. 2002. *Models of Telecenters in Brazil*. Available: http://ip.cals.cornell.edu/commdev/documents/jdc-kyle.doc (Accessed 16 February 2006).

Mbeki, T. (ed.) 1998. *Africa: the time has come: selected speeches*. Cape Town: Tafelberg.

Merriam, S. 1988. *Case study research in education: a qualitative approach*. San Francisco, Calif.: Jossey-Bass.

18

Meyer, M. 2000. How to apply the anthropological observation technique of participatory observation to knowledge acquisition for the expert system. *IEEE Transactions on Systems, Man and Cybernetics* 22(5):983-991.

Michel, J. 1997. Policy and partnerships for the development of the Infoculture. *International Information and Library Review* 24(3):465-478.

Mphahlele, M & Maepa, M. 2003. Critical success factors in telecentre sustainability: a case study of six telecentres in the Limpopo Province. Paper presented at the 2003 LIASA Annual Conference, Port Elizabeth, 3 October.

Myers, M & Young, L. 1997. Hidden agendas, power and managerial assumptions in information systems development: an ethnographic study. *Information Technology & People* 10(3):224-240.

Naidoo, J. 1997. *The role and vision of government in the emerging information society.* Speech delivered at the National Information Technology Forum Dinner. Available: http://www.doc.gov.za/docs/sp/1997/sp970801.html (Accessed 4 June 2003).

Peled, A. 2001. Outsourcing and political power: bureaucrats, consultants, vendors and public information technology. *Public Personnel Management* 30(4):26-34.

Reason, P & Bradbury, H (eds). 2001. *Handbook of action research: participative inquiry and practice*. London: Sage.

Sabien, B. 2002. *Some principles of financial sustainability based on telecentres and a telecentre network in Australia.* Available: http://ip.cals.cornell.edu/commdev/documents/jdc-sabien.doc (Accessed 17 February 2006).

Sassen, S., 2002. *Global networks: linked cities*. London: Routledge.

Schoen, R. 2002. Fitting projects to people or people to projects?, in *Participatory communication for social change, edited by* J Servaes, JL Jacobson, JL & SA White. New Delhi: Sage.

Snyman, M & Holmner, M. 2006. *The centre approach as a way to bridge the digital divide on a technological level in South Africa*. Paper delivered at the XVIth ISA World Congress of Sociology, 25-28 July, Durban, South Africa.

Snyman, M & Snyman, ME. 2001. The potential role of Information Resource Centres to create a true, open and democratic society in South Africa. Proceedings of the Open 2001 Bobbccats Conference, Vilnius University, Lithuania.

Snyman, M & Snyman, R. 2003. Getting information to disadvantaged rural communities: the centre approach. *South African Journal for Library and Information Science* 69(2):95-107.

Stilwell, P, Leach, A, & Burton, S. 1999. *Knowledge, information and development: an African perspective*. Pietermaritzburg: University of Natal.

Stones, L. 2007. Rescue attempt on poorly performing telecentres. *Myadsl.co.za,* January. Available: ttp://www.my (Accessed 7 February 2007).

Thamizoli, P & Balasubramanian, K. 2001. *Information Management and Knowledge empowerment: MSSRF Telecenters in South India*. Available: http://ip.cals.cornell.edu/commdev/documents/jdc-thamizoli.doc (Accessed 15 February 2006).

Van Audenhove, L. 1999. South Africa's information society policy: an overview. *Communicatio* 25:15-27.

Whelan, J. 2002. *Final Report: World Summit 2002*, Johannesburg. Available: http://www.iccwbo.

org/home/environment_and_energy/sdcharter/topics/sustain/VE%2520report.pdf (Accessed on 10 March 2004).

World Bank. 1998. *World Bank report.* Washington DC: World Bank.

World summit on the Information Society. 2003. *Plan of Action*. Available: http://www.itu.int/dms_pub/itu-s/md/03/ (Accessed 21 August 2005).

Yin, R. 1994. *Case study research:design and methods.* 2nd edition. Thousand Oaks, Calif.: Sage.

20

DEVELOPING A MORPHOLOGICAL NORMA-LISER FOR AFRIKAANS FOR USE IN CROSS LANGUAGE INFORMATION RETRIEVAL

ERICA COSIJN

Department of Information Science, University of Pretoria, Pretoria, South Africa
erica.cosijn@up.ac.za

ABSTRACT

This article describes the development of a morphological normaliser for Cross Language Information Retrieval (CLIR) purposes. Word form normalisation in CLIR is necessary, because plural forms, past tense verbs, etc are not included as dictionary entries and therefore cannot be translated. In the development of this normaliser, a corpus consisting of newspaper text was used to establish rules based on statistics of word form occurrences and to create a stopword list for Afrikaans. The procedure described here can normalise the majority of Afrikaans words (most past tense verb forms, most plural forms and compounds). The normaliser was tested on the original newspaper text in a CLIR environment.

DEVELOPING A MORPHOLOGICAL NORMALISER FOR AFRIKAANS FOR USE IN CROSS LANGUAGE INFORMATION RETRIEVAL

1   INTRODUCTION

The basic idea of Cross-Language Information Retrieval (CLIR) is to bridge the language boundary by providing access in one language (the source language) to documents written in another language (the target language), by using query translation from the source language into the target language and/or document translation from the target language. The main strategies for query translation are based on three different methods: dictionary-based methods with specific relevance to (bilingual) translation dictionaries, corpus-based methods, and machine translation (Oard & Diekema 1998; Pirkola et al 2001).

This study utilises query translation by means of a bilingual translation dictionary to translate Afrikaans (source language) queries into English (target language), and then matches the latter to an English language database, with English language documents as the retrieved results.

The procedure is briefly as follows: an Afrikaans query is translated into an English query which, in turn, is matched to an English language database. The result is a set of English language documents. However, in the process of translating the query from Afrikaans to English, the individual words have to be normalised before they can be translated. Word form normalisation is necessary because not all words used in everyday speech are found in a dictionary, and therefore cannot be translated. Past tense verbs, plural forms of nouns, proper nouns, complex compounds, etc are all examples of words that are not found in Afrikaans language dictionaries.

This paper specifically deals with the development and testing of a morphological normaliser so that Afrikaans words in the natural language query can be translated into English through a bilingual translation dictionary.

2   Methodology and analysis

The basis for creating the normaliser was newspaper articles in Afrikaans downloaded from various websites. The average length of the articles was about 100 words. The articles were combined to create a text corpus consisting of about 3 500 words in total. After removal of duplicates and stopwords, 1 072 unique words were used as basis for the compilation of the normaliser. Although this is not a very large sample, the normaliser was tested in a standard CLIR environment using an Afrikaans-English translation dictionary consisting of about 82 000 unique entries.

2.1 Creating a stopword list

Stopwords are words that occur very frequently and are not used when retrieving information (Korfhage 1997). Typical stopwords are "the", "an", "me", etc. The stopword list for Afrikaans was created by combining two basic methods:

- The corpus described above was sorted alphabetically and the frequency of each word occurrence was calculated. During this part of the data collection, capital letters (eg proper nouns and beginning of sentences) were ignored. As expected, the words appearing more often in the text were words that are typically regarded as stopwords.
- The top ranking words were then manually checked and compared with an existing English language stopword list. Some of the words on the English list did not appear in the Afrikaans ranked list and were added to the Afrikaans stopword list.

An interesting phenomenon was the high occurrence of the words *van*, *oor* and *deur*. These are the Afrikaans equivalent of the English prepositions "of", "over" and "through". However, in Afrikaans they also can mean "surname", "ear" and "door"

respectively, and it was initially decided to exclude these words from the stopword list because they could be meaningful words in translation. However, after the first CLIR tests, it was found that they occur many more times as stopwords than as meaningful words, and that their removal from the stopwords list had a significant negative impact on the results. It was therefore decided to include them as stopwords for future research. The current stopword list consists of 243 words.

2.2 Identifying headwords

Once the stopwords were removed from the ranked list created from the corpus, all duplicates were removed. Each of the remaining words was then compared to the Afrikaans entries in the translation dictionary we were going to use for the CLIR experiments. If the word was found "as is" in the dictionary, it was designated headword status. This meant that these words could be translated without any normalisation processes being applied to them. This process was partly automated, but some manual input was necessary to distinguish between proper nouns starting with capital letters, and other words starting with capital letters because they were the first words in a sentence. Proper nouns are typically not found as dictionary entries, whereas the other words starting with capital letters had to be rewritten with small letters to check if the word appeared in the dictionary as a headword entry.

Once the stopwords were removed and the headwords were identified, a normaliser had to be created to deal with all the remaining words. The remaining words were typically of the following types: plurals of nouns, compounds, proper nouns, past tense forms of verbs, adjectives formed from verbs, etc. Another complication was the relatively high occurrence of incidental diacritical marks in Afrikaans (see section 2.9 below).

The remaining words were now manually analyzed and categorised, not according to the rules of spelling or grammar, but according to statistical patterns, to create algorithms that would normalise the words to headwords through rules. Sections 2.3 to 2.7 describe the categories identified.

2.3 Double vowel cases

Most nouns in Afrikaans containing *-aa-*, *-ee-*, *-oo-* or *-uu-* in the last syllable of the singular form, drop one of the double vowels and an *-e* is added to the end of the plural form. This is also the case in the formation of some adverbs and adjectives. These words were manually identified from the list of remaining words. Table 1 shows some examples of "double vowel cases".

The rule created here was that, if a word cannot be identified as a stopword or a headword, the algorithm should check whether the last three letters of the word is any of the following: *-a-consonant-e*, *-e-consonant-e*, *-o-consonant-e* or *-u-consonant-e*, in

23

which case the *-e* is dropped and the vowel in front of the consonant is doubled. If the normalised form is found as an entry in the dictionary, the word is translated.

Table 1: Some examples of successful normalisation of "double vowel cases"

Word in test corpus	Dictionary entry	Comment
Palmbome	*Palmboom*	*boom* is the singular form of *bome*
Seksuele	*Seksueel*	*seksuele* is an adjective formed from *seksueel*
atlete	*Atleet*	*atlete* is the plural form of *atleet*

2.4 Suffix stripping

The remainder of the corpus was now manually checked for all word forms containing suffixes or, more precisely, for all word forms consisting of a headword entry and appended with any number of letters, but excluding compound words (Korfhage 1997). The idea was to find the most common suffixes in Afrikaans and write a rule to strip these from the headword, so that it is possible to find it in the translation dictionary.

During the analysis it was found that there are quite a number of suffixes and therefore it was decided to split the process into two stages. The first stage was to identify only nouns where the plural is formed by adding an *-e*, *-s* or *-'s* to the singular form. The second stage was to identify nouns, as well as other word forms where the headword is appended by any type of suffix.

During the first stage it was found that almost 80% of all plurals in the test corpus was formed by adding the suffixes *-e*, *-s* or *-'s* to the singular noun. These plural forms constituted 7,9% of the total number of words in the test corpus. (Also see Table 3.)

The second stage identified *-te*, *-ste*, *-er*, *-se*, *-der*, *-de* and *-ing* as the most common suffixes in the test collection. However, after some manual checking we realised that simply assuming that these strings are suffixes to be removed might have serious adverse effects on the translation process. These endings to words were found not to be suffixes in all cases, but were sometimes also real parts of word endings; for example, *roete* (route), *passasier* (passenger), etc.

The same problem would have occurred in the first stage analysis, where many headwords end in *-e* or *-s*, and where automatically stripping these suffixes from words would have led to mistranslations. The possibility of longest common substring (LCS) matching was therefore investigated.

2.5 Longest common substring matching

LCS matching means that, when two character strings are compared, the longest continuous matching string of characters is identified (Gusfield 1999); for example, the

LCS of "antimatter" and "material" is "mat".

Through manual analysis of all the words identified in the two stages described under section 2.4 above, it was found that a much more desirable outcome would be reached if, instead of suffix stripping, longest common substring (LCS) matching was used. Afrikaans is not an inflectional language and, according to an analysis of the test corpus, prefixes are not often used (with the exception of the past tense prefix *ge-*, which is discussed in section 2.6 below). It was thus assumed that, with the exception of words starting with *ge-*, words in normal Afrikaans text can be matched to dictionary entries by using LCS matching from the beginning of the word, starting with the first three letters of the word. This would successfully normalise all possible words containing suffixes, including the plural forms created by adding an -*e*, -*s* or -*'s* to the end of the noun.

The process works as follows: the word is matched from the beginning, starting with the first three letters and progressively longer strings are tested for a match. When a match is found, the process does not stop, but is continued until the longest possible matching string between the word and the dictionary entry is matched. As an example, *swembaddens* is matched as follows:

swe	no match
swem	matches *swem* = "swim" in the dictionary
swemb	no match
swemba	no match
swembad	matches *swembad* = "swimming pool" in the dictionary
swembadd	no match
swembadde	no match
swembadden	no match
swembaddens	no match

The Longest Common Subsequence (LCS) match is thus *swembad* = swimming pool.

Table 2 shows some successful matches following this procedure.

Table 2: Some examples of successful normalisation with LCS matching

Word in test corpus	Dictionary entry	Comment
swembaddens	*swembad*	Plural form where suffix -*dens* is added to singular
permanente	*permanent*	Adjective formed where -*e* is added to an adverb
Verteenwoordigende	*verteenwoordig*	Adjective formed where suffix -*ende* is added to a verb
boetes	*boete*	Plural form where suffix -*s* is added to singular

25

2.6 Prefix stripping

The only prefix found in significant numbers in the test collection was *ge-*, which is a prefix added to a verb to indicate the past tense or, sometimes, as an auxiliary verb to indicate a continuous action. The rule created to deal with this class of words was thus: if the word cannot be matched to the dictionary entry and it starts with *ge-*, then remove the prefix and attempt to match again. If there is a match, then translate the word. If there is no match, then the word is retained and taken to the next step in the process. For example: *gedoen* ("did") does not match any entry in the dictionary. It starts with *ge-*, therefore drop the prefix and match *doen*. *Doen* matches "do" in the dictionary, and is thus translated correctly.

2.7 Compound splitting and fogemorpheme identification

As in most Germanic languages, compounds are quite common in Afrikaans. It was calculated that almost 6% of all the words in the corpus were compound words. In Afrikaans, compounds may be formed "on the fly" (Carstens, 2003) and not many compounds are included as dictionary entries (in this dataset, 6% of the total number of words were compounds: 1,3% were found in the dictionary and 4,7% were not). In order to translate compounds, therefore, they first have to be split into their component parts and these then have to be translated separately.

From the compounds in the test corpus, it was found these are either simply formed by joining two or more words, but that in some cases a fogemorpheme (joining morpheme) is added to aid pronunciation. Two fogemorphemes were identified, namely -*s*- and -*e*-.

It was decided to also use LCS matching to split and translate the compound. If a joining morpheme is included in the compound, there has to be an extra step in the process: once the longest common subsequence has been established from the front of the word, and the second part of the compound cannot be matched, then the algorithm should determine whether or not the next string starts with either an *e*- or an *s*- and, if it does, this letter should be discarded and the matching started again from the second letter of the rest of the word. For example, the word *kliëntekontaksentrum* ("client contact centre") contains a fogemorpheme -*e*- after *kliënt*- and can be split as follows:

kli	No match
. . .	
kliënt	Matches *kliënt* = "client" in the dictionary
kliënte	No match
kliëntek	No match
. . .	
kliëntekontaksentrum	No match

The LCS is thus *kliënt*. According to the procedure, *ekontaksentrum* should now be matched. Applying LCS matching from the first three letters, *eko..*, *ekon...* etc no matches are found. Because there is no match, a second LCS matching procedure is initiated. In this second step the *e-* is dropped and *kontaksentrum* is matched thus:

Kon	Stopword
. . .	
Kontak	matches *kontak* = "contact" in dictionary
. . .	
Kontaksentrum	no match

LCS is thus *kontak*. The matching process starts again with remainder of the word, thus:

Sen	no match
. . .	
Sentrum	matches *sentrum* = "centre" in dictionary

klientekontaksentrum is thus matched to *kliënt + kontak + sentrum*; each of these words is then translated separately and the full compound is correctly translated as "client contact centre".

After the creation of the rules described in sections 2.3-2.7 above, all the unique words (N=1 072) were manually categorised to establish what percentage are translatable or normalisable. See Table 3 for details.

Table 3: Stopwords, headwords and words that can be normalised

Types of word	N	%
Stopwords	150	14,0
Headwords	565	52,7
Double vowel rule solvable	18	1,7
LCS matching to solve plurals of form -*e*, -*s*, -*'s*	85	7,9
LCS matching to substitute suffix stripping	59	5,5
Stripping prefix *ge*-	13	1,2
Compound splitting (multiple LCS runs + fogemorpheme stripping)	50	4,7
Total	**940**	**87,7**

2.8 Not solved

After removal of stopwords through the stopword list and then matching the words that appear in the same form as the headword in the dictionary and the normalising processes described above, 7,5% of the total list were left as proper nouns not in the dictionary and 4,8% of the total list could not be solved by any of the procedures described above. The make-up of the 4,8% unsolvable words is shown in Table 4.

27

Table 4: Words that could not be normalised using the rules created

Type	N	%	Examples
Compounds incorrectly solved	8	0,7%	*Eenrigtingpaaie* (one-way roads) splits into *eenrigting* + *paai*. The second part of the compound is plural for *pad* (road), which cannot be normalised by the rules created. However, *paai* is a headword in the dictionary meaning "appease" and the entire compound is mistranslated.
Past tense -*ge*- embedded in word	26	2,4%	opgerig nagegaan afgelewer
Not solvable	16	1,5%	New words, eg CD, DVD Words not in the dictionary Rare, archaic and other word forms that cannot be identified by LCS matching (eg nuwe --- adjective from nuut [new])
Misspellings	2	0,2%	Misspelled words in the newspaper text
	52	4,8%	

2.9 Incidental diacritical marks

Diacritical marks are generally quite common in Afrikaans; for example, *sê* ("say"), *rûe* ("backs"), *bôgom* (the noise a baboon makes), *priële* ("pergolas"), *weeïg* ("faint"), *séance* ("séance"). These words are in the dictionary as such, and an expanded character set (ISO 8859-3) had to be used in the programming of this normaliser.

An interesting phenomenon, unique to Afrikaans, is the addition of diacritical marks (specifically the acute mark) whenever words are to be emphasised. For example:

Dáár is 'n perd --- THERE is a horse (meaning very far away) as opposed to

Daar is 'n pérd --- There is a HORSE (meaning not a cow).

These acute markers may be placed on any vowel in the word or syllable that is to be emphasised (with a few exceptions).

The problem encountered in automatic text analysis is that these accented words are not found in the dictionary. Possible ways of dealing with this problem are as follows:
- Duplicate all words in the dictionary that may conceivably have acute markers (this would involve many hours of manual input).
- Replace all letters with diacritical marks with the same letter, but without the diacritical mark (this would have a negative impact on words that have diacritical marks as part of their correct spelling).
- Leave these words as untranslatable.

Analysis of the test data showed that stopwords (especially articles and negations) are

28

often emphasised, but nouns and verbs less often. It was therefore decided to duplicate all the stopwords (with and without acute signs), but leave other words with incidental diacritical marks as untranslatable.

3 Morphological normaliser

After the manual analysis of the test data, it transpired that a large percentage (nearly 90% --- see Table 3) of words in Afrikaans are stopwords, uninflected words, or solvable by one or more of the rules described under sections 2.3 to 2.7.

The next step was to establish the most sensible order in which the rules should be applied. This is a very important consideration, because once a word is matched, it is translated and cannot be processed again. Figure 1 (originally drawn by Heikki Keskustalo) illustrates the process described below.

After several trial runs, the steps were listed as follows:

Step 1: Remove all stopwords through comparison with the stopword list.

Step 2: Match all remaining words with headwords in the dictionary. If a match is made, the word is translated.

Step 3: Of the remaining words, check if a word starts with a capital letter --- if it does, it should be replaced with a small letter. This includes proper nouns and all words at the beginning of sentences. The reason for including this step is that, if a word starts with a capital letter, it is not found in the dictionary. None of the words that appear at the beginning of a sentence, therefore, will be translated. If the word is a proper noun, the assumption is that it will not be found in the dictionary in any case, and will be fuzzy matched in the last step (step 9).

Step 4: Of the remaining words, check which of these contain hyphens (hyphenated compounds). The hyphens should be removed and replaced with a space and the individual words then checked against the dictionary. (Note: some compounds are hyphenated, but some headwords have hyphens too. Headwords should therefore be checked first (step 2) before the removal of any hyphens in order to match those dictionary entries that are, in fact, hyphenated.)

Step 5: Identifying simple past tense verbs: if the word starts with *ge-*, remove the prefix and check the remainder of the words against dictionary headwords. If a match is found, the word is translated.

Step 6: Plurals are now identified where, if the last three letters of the word are a single *o*, *a*, *e* or *u* followed by one consonant and an *e*, the last *e* is stripped and the vowel before the consonant is doubled.

29

Step 7: Longest common subsequence matching is utilised to translate words with suffixes (as identified in section 2.4), as well as plurals ending in -*e*, -*s* or -*'s*.

Step 8: Matching compounds that are not headwords in the dictionary require multiple LCS matches. The procedure is to start from the beginning of the word and match the longest possible headword. Repeat the process until no more matches can be made. If a match cannot be made, check if the next section of the word starts with an *e* or an *s*, remove and check against the dictionary again.

Step 9: All the remaining words are now capitalised, because it is assumed that they are probably a proper noun (proper nouns accounted for about 7,5% of the total number of unique words in the test data). These are not translated, but fuzzy matched directly to the English database. For example, *Pretoria* is reduced to *pretoria* in step 3. It is not translated by any of steps 4-8, and in step 9 therefore becomes *Pretoria* (Afrikaans) again and is matched with Pretoria (English) in the database.
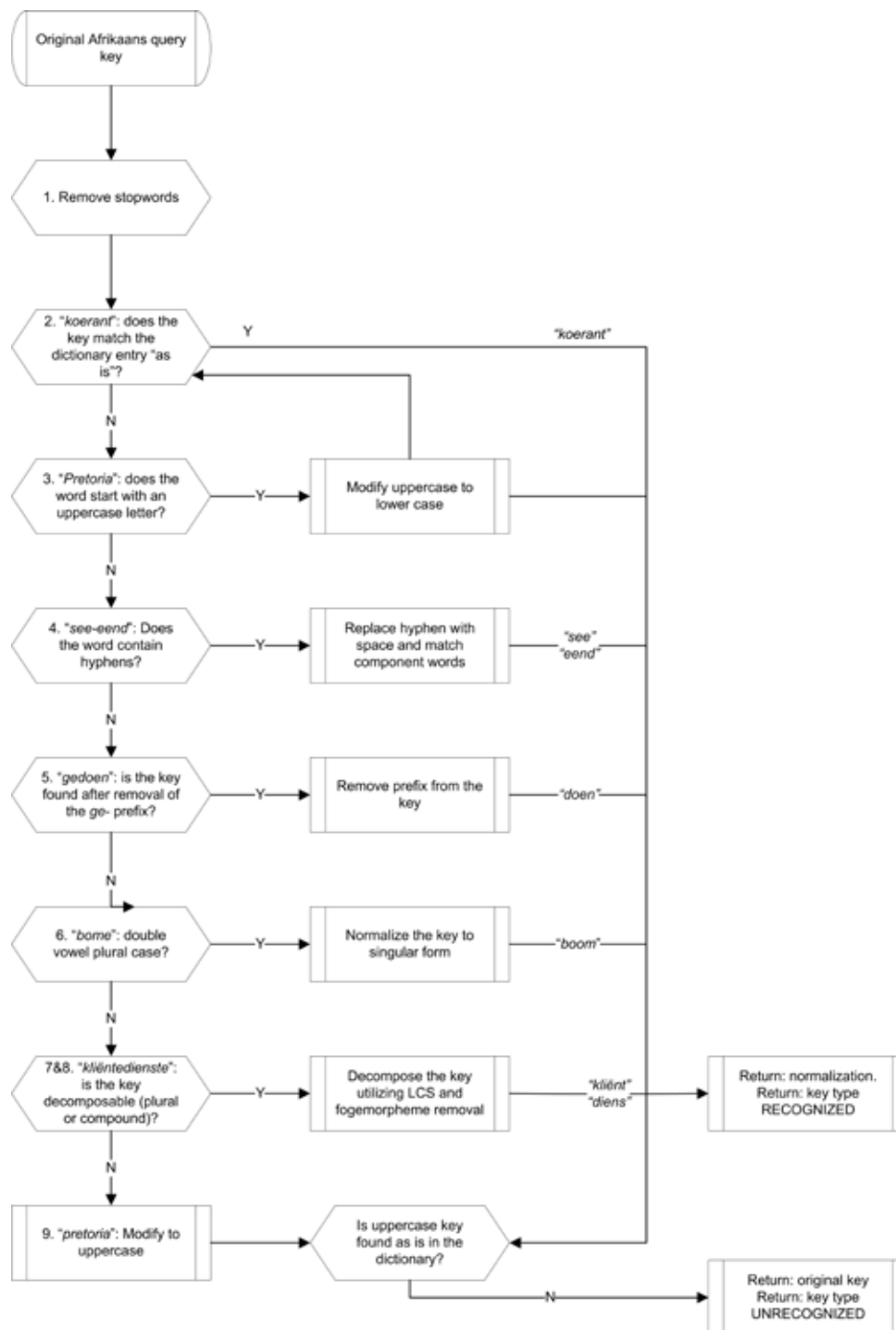
Figure 1: Afrikaans normaliser

31

4   Results of testing

4.1 Testing the original dataset

After the normaliser was programmed, the same dataset was used for testing. Table 5 shows a random example of the results. The first column shows the words in the original Afrikaans text, the second shows the automatic translation through the dictionary (consisting of about 82 000 unique headwords), and the last column contains comments on the translation after a manual analysis of the results.

Table 5: A random sample of translations of the test data

Afrikaans word in newspaper text	Automatic translation after normalisation	Comment
adjunk-direkteur-generaal	adjunct deputy assistant director-general	Compound split and translated correctly after removal of hyphens.
adjunkdirekteur-generaal	adjunct deputy assistant director-general	Compound split and translated correctly. Note the different ways of spelling (see row above): both forms translated correctly.
ag	eight attention care esteem deem respect value necessary oh ah alas	Found in dictionary as is.
agtereenvolgende	consecutive successive	Found in dictionary as is.
gedompel	plunge dive dip immerse	Removal of *ge-* prefix. Remainder of word found in dictionary and translated correctly.
Standaardgraad-aardrykskundevraestel	standard criterion archetype gauge grade degree rate stage grade gradation geography examination paper	Capital initial letter replaced with small letter, thereafter compound split and translated correctly. This compound consists of four terms: *standaard* + *graad* + *aardrykskunde* + *vraestel*. Correctly split and translated.
bome	tree bar beam	Double vowel plural case: *bome* did not match, the last three letters were identified as *o-consonant-e*; it was then reconstituted to *boom* by dropping the -*e* and doubling the vowel in front of the consonant. The resultant word was then correctly translated.
rede	sense reason account	Correctly translated. This was not recognised as a "double vowel plural case", since the order of translation is to match words first as is.

32

The manual analysis of the translation of the words in the original dataset did not show any significant problems regarding the morphological normaliser.

4.2 Testing in a CLIR environment

The detail concerning the testing, methodology and results of the first test for Afrikaans-English Cross Language Information Retrieval can be found in Cosijn et al (2004). A summary of the experimental setup and results follows below.

The test system was the InQuery retrieval system (Allan et al 2000). InQuery is a probabilistic retrieval system based on the Bayesian inference net model. Queries can be presented as a bag-of-words queries, or they can be structured using a variety of query operators. All query keys are attached with a belief value, which is calculated by the tf.idf modification of the system.

The test data of this study was CLEF 2001 data (Peters 2002). The data include a test collection and a set of test topics. The collection contains some 112 000 newspaper articles from the *Los Angeles Times*. The test topic set used in this study contained 35 of the 50 CLEF 2001 topics, and both the short form (title) and long form (title and description) were tested. The baseline for comparison is English queries to English database. The Afrikaans to English query translation is based on the UTACLIR (University of Tampere Cross- Language Information Retrieval) framework described in Hedlund et al (2004). (Also see Hedlund et al (2004) for a discussion on UTACLIR and CLEF 2001.)

The results (Cosijn at al 2004) were evaluated as (1) average precision over ten recall points (10-100%) and (2) as precision at the 10% recall point. The results are presented in Table 6.

Table 6: Findings of retrieval experiments

Query Type N=35	Average Precision (%)	% of Baseline	Precision at 10% Recall	% of Baseline
Title				
Baseline	32,6		60,1	
Test queries	13,6	37,6	25,0	41,6
Title + description				
Baseline	39,8		63,1	
Test queries	19,4	48,7	32,8	52,0

As shown, the average precision of the test queries was 13.6% (title only) and 19.4% (title and description), while baseline queries give an average precision of 36.2% (title only) and 39.8% (title and description).

The relative performance of test queries with respect to the baseline queries was 37.6% (title only) and 48.7% (title and description). At the 10% recall point, test queries performed better with respect to baseline queries, with the relative performance figures being 41.6% (title only) and 52.0% (title and description).

The main problems that were experienced related to the following:
- Untranslatable words (words not found in the dictionary, proper nouns, abbreviations and acronyms, adjectives formed from names of countries and rare inflectional forms that are not solvable by the normaliser).
- Compounds where part of the compound is not in the dictionary (especially computer related terms and proper nouns).
- Homonyms and lexical ambiguity, especially in the case of words that are stopwords, but also have a meaningful sense (eg oor, van and deur as described in section 2.1 above).
- Afrikaans set expressions as part of the dictionary entry with an English translation attached. The ideal would be to have a simplified version where there is only the Afrikaans headword and the English translation(s) of the headword.
- Incidental diacritical marks were relatively easily solved for stopwords (see section 2.9 above), but where these occurred in the running text of the query, the words were mistranslated because they did not match any dictionary entries.

5 Conclusion and recommendations

In general, the results were quite promising and the normaliser, although relatively simple, was found to perform quite well. The problems experienced above are not necessarily due to the morphological normaliser itself.

The limitations that are specifically related to the morphological normaliser include the following:
- The normaliser cannot process plurals that are not double vowel cases, or where the plural is formed by adding an -*e*, -*s* or -*'s*. Examples are: *skip – skepe*, *pad – paaie*. These two specific examples will not be solved by LCS matching either, but many other cases can be solved by LCS matching (eg *bad – baddens*, *gewas – gewasse*, *lem-lemme*).
- The normaliser cannot process words where an affix is embedded in the word. This is often the case in the past tense of a verb that is combined with a preposition (eg *aangekom*). In this case, *aan* is the preposition; *kom* is the verb, and the -*ge*- indicating the past tense is embedded in the middle of the word. A possible solution might be to include a step between steps 8 and 9 (section 3) that will strip stopwords as prefixes (all prepositions are typically stopwords) if a stopword+ge

34

combination is encountered. Once the preposition and -*ge*- is removed, the remainder of the word can be matched to the dictionary.

On the basis of the tests with the original as well as the CLIR experimental dataset, the performance of the morphological normaliser was not the main problem where mistranslations occurred. It is likely that, with a little fine tuning to overcome the problems and limitations stated above, the normaliser will perform optimally within an information retrieval environment.

Acknowledgements

References

Allan, J, Connell, ME, Croft, WB, Feng, F-F, Fisher, D & Li, X. 2000. *Inquery and TREC-9. The Ninth Text REtrieval Conference (TREC-9), Gaithesburg, MD.* Available: http://trec.nist.gov/pubs/ trec9/t9_proceedings.html (Accessed 3 May 2007).

Carstens, WAM. 2003. *Norme vir Afrikaans: enkele riglyne by die gebruik van Afrikaans.* 4de uitg. Pretoria: Van Schaik.

Cosijn, E, Keskustalo, H, Pirkola, A, De Wet, K & Järvelin, K. 2004. Afrikaans-English cross-language information retrieval, in *Progress in library and information science in Southern Africa. Proceedings of the third biennial DISSAnet conference (ProLISSA3)*, edited by T Bothma & A Kaniki. Pretoria: Infuse: 97-110. Available: http://www.dissanet.com/jsp/modules/repository/index.jsp?repository=Library&action=view_file&id=Library.1109248798718 (Accessed 27 June 2007).

Gusfield, D. 1999. *Algorithms on strings, trees and sequences: computer science and computational biology.* Cambridge: University Press.

Hedlund, T, Airio, E, Keskustalo, H, Lehtokangas, R, Pirkola, A & Järvelin, K. 2004. Dictonary-based cross-language information retrieval: learning experiences from CLEF 2000-2002. *Information Retrieval* 7(1/2):99-119.

Korfhage, RR. 1997. *Information storage and retrieval.* New York: Wiley.

Oard, D & Diekema, A. 1998. Cross language information retrieval. *Annual Review of Information Science and Technology* 33:223-256.

Peters, C. 2002. *CLEF – Cross-Language Evaluation Forum.* Available: http://clef.iei.pi.cnr.it:2002/ (Accessed 2 June 2007).

35

Pirkola, A, Hedlund, T, Keskustalo, H & Jarvelin, K. 2001. Dictionary-based cross-language information retrieval: problems, methods and research findings. *Information Retrieval* 4(3/4):209-230.

Relevance assessment for digital libraries

Jela Steinerová

Department of Library and Information Science, Comenius University, Gonsova 2, 818 01 Bratislava, Slovak Republic
steinerova@fphil.uniba.sk

Abstract

This article discusses a study carried out on the relevance of information use as perceived and experienced by doctoral students. Relevance is studied with the use of phenomenographic methodology. Twenty-one students were interviewed with the aim of interpreting their relevance judgments. Analyses are presented in terms of the potential for new patterns of relevance for digital libraries research. Selected results are interpreted from the viewpoint of perception, emotions, metaphors of relevance, criteria, strategies in electronic environment, and organisation of information. In perceptions of relevance, ideas of importance, utility, and value prevail. Metaphors refer to cleaning from the non-substantial, surprise, and understanding as a spontaneous experience. Multiple criteria are applied intuitively as part of situated actions. Emerging patterns of relevance in the electronic environment indicate the need for flexibility, interactivity and collaborative use. Hierarchy and associations of knowledge organisation support relevance judgments. At the same time, creativity and discovery should be allowed. The final conceptual model of the perception and experience of relevance points to a need for participation and decision-making frameworks. Backtracking, serendipity and interaction between individual and collective spaces can add value to relevance patterns for digital library use.

Keywords

- relevance
- phenomenography
- interviews
- information use
- digital libraries

1. Introduction

Only limited attention has been paid to the use of information in digital libraries. Relevance was usually studied specifically in terms of an evaluation of services, system performance, or the success of the digital library. In this research, we shall try to discover new aspects of relevance by applying phenomenological background. This research is based on the assumption that relevance is a specific experience in information use. The naturalistic paradigm determines human beings in contexts and applies qualitative methods. This paradigm is convenient for studies of human beings, their behaviour and experiences. We intend to apply qualitative methods in order to investigate relevance judgments with implications for digital libraries. In this type of research, we assume that relevance is not a computation or an algorithmic matching, but instead a way of communication and an everyday experience.

The concept of relevance can be interpreted as links between people and information directed to creating meanings. Some of the links become borders which determine what will, and what will not be, included into the cognitive structure. One way to improve the success of digital libraries could be a phenomenological inquiry into the experience of users when judging relevance.

In this paper we point to the benefits of links between phenomenological relevance studies and digital libraries research. We report on the findings of a study of doctoral students with respect to relevance judgments. PhD students represent a special group of digital library users, a group that integrates both education and research. Their patterns of relevance judgments can be used by digital libraries developers for designing value-added services. A special perspective of the research is its setting in central Europe.

The rest of the paper is organised as follows. Section 2 presents related work as a background of the study. Further sections concentrate on the case study based on our own research of relevance, including goals, methods and results. Section 6 concludes the paper with a synthesising conceptual model derived from analyses of the empirical material.

2 Related work

A successful digital library was explained by several models (Seddon, DeLone, Venkatesh), including a new digital library success model (Shen, Vemuri, Fan & Fox 2006: 208-219). Another model is represented by a framework of information seeking and retrieval in context (Ingwersen & Järvelin 2005: 261). These models synthesise information seeking models (eg Ellis 2005: 138, Kuhltau 2005a) with information retrieval and evaluation of digital libraries. User behaviour in digital libraries is exemplified

by sophisticated interface interactions. Evaluation of relevance is used in knowledge augmentation manifested by digital library interactions (eg discussions, annotations, comments, etc). (Frommholz & Fuhr 2006: 279-290).

In line with several synthetic works on relevance (Froehlich 1994: 124-134 ; Saracevic 1996: 201-218 ; Mizzaro 1997: 810-832) we do not make an attempt to define relevance. Instead, we are interested in understanding the phenomenon of relevance by means of natural human information behaviour. Naturalistic approach in information behaviour research is represented especially by Dervin´s sense-making methodology (Dervin 2003). Other examples include the socio-cognitive relevance (Hjørland 2002: 960-965), the user-centred relevance (Bruce 1994: 142-148) and the discourse analysis (Talja 2005: 123-127). Our assumption is that manifold information needs and various situations of information use give rise to manifold relevancies in different contexts. The relevance judgment patterns can be mapped in models of domain knowledge, work tasks, granularities of knowledge organisation, strategies, information needs and types, contexts, motivations, and constraints.

Similar studies concentrated on understanding users´ experiences in order to evaluate digital libraries (Blandford 2003). Relevance is implicit in the large evaluating frameworks of digital libraries (eg Borgman and Larsen (2003) and Saracevic (2004).

We believe that relevance in digital libraries is driven by cognitive and interactive foundations of information use. Marchionini (1997) referred to information use from the perspective of information seeking in electronic environments. The principle is information extraction mediated by conceptual information handling, especially reading, classifying, storing information, etc. A networked information environment triggers new patterns of relevance assessments as part of information use. New patterns of relevance are required for digital libraries and new web services. These patterns apply cognitive and interactive principles of information use and can enhance personal information management, personal experiences and collaboration. Concepts of situational and psychological relevance (Harter 1992: 602-615), but also graded relevance (Maglaughlin & Sonnenwald 2002: 327-342), (Spink & Greisdorf & Bateman 1998: 599-621) confirm functions of multiple criteria, decision frameworks, situatedness and contexts of information use.

Recent relevance studies emphasise relevance in action and make use of qualitative methods (Anderson 2006). Relevance is interpreted as an experience, connections and communications. Qualitative methods can disclose implicit characteristics of relevance emerging from experience, preferences, collaboration. In the networked environment, more value and contexts can be added to information by conversation, sharing, categorisation, recommendation, visualisation.

Bates´s berrypicking model (Bates 2005: 58-62) contributes to new perspectives of relevance in the web environment. Interactions, changes of knowledge states and different strategies of information seeking become driving forces of relevance judgments. Non-linear, exploratory and intuitive characteristics of information seeking are mapped in relevance judgments. The question is how interactive interconnections, enriched interactions and collaborative uses lead to relevant information.

3 Case study: objectives and questions

Based on these assumptions, we designed an empirical study of PhD students´ information behaviour in relevance judgments. The study was aimed at discovering perceptions of relevance and deriving proposals for more effective ways of information use. The methodological background draws on naturalistic-phenomenographic and constructivist concepts. We believe that social methods of human information behaviour can help discover new perspectives on digital libraries research.

Principal research questions were articulated as follows:
- What are the perceptions of relevance with PhD students from various disciplines?
- In what ways is relevance experienced in terms of applied criteria, subjective and objective components?
- Which categorisation, types of information and what factors influence relevance judgments?
- How is relevance manifested in the electronic information environment?

Data were acquired using semi-structured interviews with 21 PhD students and master graduates of the Faculty of Philosophy, Comenius University Bratislava, Slovakia. Participants were selected from various disciplines of social sciences, the majority of them having a library and information science background. With PhD students, the information use in the areas of science and education is integrated. All of them have real information needs and previous experience in researching a topic. Indeed, this is why they were a convenient research group. Our choice of students was based on recommendations from their supervisors.

The structure of the interviews was designed as a set of open questions. We tried to cover the complexity of the investigated phenomenon, drawing on related studies of relevance (Anderson 2005) and phenomenography (Limberg 2000: 51-68). Appendix 1 lists the questions used in the directed, semi-structured interview. The goal was to capture the perception and interpretation of relevance issues, and follow one's own, original verbalisation of relevance.

Instructions for the guidance of interviews explained the goals and contexts of the research in detail. The interviews were conducted by two researchers over five months (October 2005-February 2006). The interviews lasted from between 25 to 60 minutes. Subjects were between 24 and 47 years of age; the average age was 28, 9 years. Gender distribution was 12 women (57 %) and nine men (43%). The participants had different amounts of practical experience, ranging from three months to more than 20 years. As for the disciplines studied, students´ majors were diverse, 13 subjects came from library and information science (61%) and eight from other disciplines (39%) (philosophy, ethnology, journalism, psychology, history).

4 Methods

The semi-structured interviews consisted of open-ended questions structured into an introductory part, and then the interview itself. Questions concentrated on the rise of the information need and previous experiences with professional information. In the first part, participants were asked to explain their idea of relevance. After this, we asked them to recall activities, decisions, and criteria which they had to apply to relevance judgments. We then examined differences in relevance judgments in the electronic and printed environment. Further questions concerned auxiliary sorting of information, influences, supportive intuitive clues, and relationships between subjective and objective components. We also asked them to identify differences between the orientation stage and the problem-solving stage. The affective component of information behaviour was represented by questions about emotions and the most interesting part of relevance judgments. The interaction between cognitive and affective components was integrated by a question on metaphor or simile.

Since the interview required deeper, abstract thinking, the participants needed more time in which to articulate their answers. Some of them found several questions similar; some said they would have preferred written questions. Others said they did not really have the time to participate; still others had problems understanding the question, or suffered from overload (eg questions too abstract). All comments can be represented by one participant's comment: "It was nice talking ..., but I feel dizzy already".

The interviews and discussions were recorded and transcribed. Records, transcripts and notes from interviews were used for analyses. Three researchers analyzed and interpreted the study data in several steps. In the first step, the answers were analyzed and coded in terms of key concepts. The answers were then integrated with regard to single questions. In the end, 21 answers were recorded.

We also used quantitative analyses, especially concept frequencies and occurrences. Content conceptual analysis concentrated on sub-issues related to single questions. The

extracted concepts were categorised by common attributes. Broader categories were derived from quantitative occurrences of semantically related concepts. In some cases, a concept map of broader categories and concepts was created (eg understanding and perception of the concept of relevance). Based on semantic analyses, semantic models were derived. In other cases, expressions were analyzed in detail by a comparison of the "pros and cons" of the problem in question (eg traditional versus electronic environments). Tables of concepts and their occurrences were elaborated.

Resulting conceptual structures were discussed and further categorised at different levels of abstraction. The validity and reliability of results are provided by three independent analyses and interpretations. Selected participants also evaluated the results. In the closing stages of the research, interpretations of collective discourse and new models emerged. The models generalise conceptual analyses and derive conclusions that not only describe participants' statements, but also explain possible contexts and causes of the issues examined.

5 Results

In this section, results of analyses are interpreted in terms of perceptions of relevance, metaphors, emotions and moods, applied criteria, relevance in the electronic environment, and organisation of information. Results also cover utilisation practices and relevance behaviour patterns. Here we have chosen those examples which introduce newer views on relevance resulting from applied phenomenographical methods.

5.1 Perception of relevance

By asking the question "In your view, what is relevance of information?", we focused on the intuitive understanding of relevance, as implicit in everyday language. The aim was to discover a perceived collective meaning of the concept. Tables and visual models were derived representing results of content analysis. See Table 1. Based on this we modelled the prototype meaning of relevance as it emerged in the collective discourse of participants. This prototype consists of the following three categories: value, utility, and importance of information.

Value of information represents its internal integrity, validity and reliability. Participants related this meaning with a process, often as part of the accomplishment of seeking (eg "way to knowledge", or "guidance by goal"). A number of them emphasised verification, authentication, and credibility of information. Trustworthiness of the source and verity of information were also regarded as significant.

Utility narrows the extension of value towards a more concrete use of information.

42

Utility is embedded in the contexts of information use, namely in relation to topic, problem solving, and time ("the most important information in a minimum of time"). Additional issues covered by utility are appropriateness and topical matching. Topical matching was expressed by appropriateness, accuracy and by question-answering (concrete context).

Importance expresses qualities of relevance such as emphasis on the problem essence, priorities, and the hierarchical division of information (important, peripheral). Participants emphasised focal, carrying information, implicit categorisation, and the subjective nature of relevance. It is connected with taking up a position and an ability to "select".

This was explicitly expressed by one student, who explained relevance as information which I need "in my state". Although traditional, established ideas of relevance prevailed, they have often been enriched with the value of a resource (reputation, verity).

Concept	Category		
	Meaning	**Context**	**Process**
Importance	6		
Value	4		
Utility	3		
Accuracy	3		
Usability	2		
Recency	2		
Feasibility	2		
Category	2		
Adequacy	1		
Truthfulness	1		
Substantiality	1		
Verified	1		
Credibility	1		
Trustworthiness	1		
Ability	1		
Concept	1		
Accuracy	1		
Quality	1		
Sufficiency	1		
Need		4	
Time		2	
Topic		2	
Demand		2	

43

Goal		2	
Solution		1	
Source		1	
Answer			1
Decreasing			1
Selection			1
Transformation			1
Position-taking			1
Formulation			1
power of category*	**35**	**14**	**6**
number of concepts within category	**19**	**7**	**6**

Table 1 Example of relevance perceptions analysis

5.2 Emotions, moods, and attractiveness of relevance assessment

We categorised participants' feelings by using basic approaches to human emotions: agreeable (positive) emotions, and disagreeable (negative) emotions. In modelling we applied the following dimensions: (1) positive/negative, (2) agreeable/disagreeable (evaluation), (3) tension/relief (activation), (4) excitement/ease (grade of intensity). Evaluation and activation are complementary. Emotions regulate relevance judgments in selecting information, associating information with knowledge in memory, and organising and recollecting information.

Content analysis showed that, at the level of agreeable feelings, the most often used expression was *delight.* This refers to a pleasant feeling of achieving something desired. Different grades of intensity of delight were manifested by concepts of satisfaction and happiness. Some participants expressed their satisfaction by metaphoric statements --- e.g. "I can check this off (my list)", or "the subjective scientific delight". *Enthusiasm, encouragement*, and *discovering* were identified as other agreeable feelings (eg "when I enjoy myself", "it is a gift to be able to read").

Negative feelings have been verbalised by higher grade of activation, namely *anger* as reactions to obstacles while achieving a goal. Contexts of anger have been explained from different perspectives; for example, anger at oneself ("that I was not able to better articulate the requirement"), or anger as rage ("when information is based on false points"). Other negative feelings were grouped around the concept of *fear,* including uncertainty, helplessness, and anxiety. This is related to one's own inability to control the situation, which is characterised by large "unknowns" ("dead ends, when I am un-

able to move forward"). Novelty and extraordinary nature of situations in relevance judgment can cause both anxiety and distress. Most frequently mentioned causes were time pressure and task overload.

Additional feelings of fatigue, resistance, and disappointment emerged in categorisation. Some participants expressed their attempts at "no emotional engagement".

However, emotions play an important part of relevance judgment. They operate in interactions with cognitive processes; feelings are integrated with the evaluation of information and are manifested at different levels of intensity. Positive feelings of relevance judgments lead to activation, vitality, self-confidence, and self-reflection. Desired information is regarded as possessing extraordinary value (a view based on an implicit value system). In science and education something new and latent is discovered that is directed towards scientific truth. Negative feelings within relevance judgments include reactions to obstacles and limitations. These feelings are based on a lack or surplus of information. The problem is how to control the situation conditioned by novelty, strangeness and surprise caused by information. Imbalances between expectations and results, as well as lack of time, are also frequently given as reasons for negative feelings.

Participants confirmed that the role of emotions in relevance judgment was more vital than anyone has ever really admitted. They experienced emotions as a discrepancy between their expectations and practices of relevance judgment. Participants reported that they felt a progress away from disagreeable feelings of uncertainty and tension to feelings of relief. However, this process is not straightforward. It is highly individual and changes with context.

Following the perception of the attractiveness of relevance judgment, we asked participants to identify the interesting components of the process. They emphasised especially *creative thought, finding a value, and discovering something new.* Some of them appreciated the experience of understanding the problem, inspiration and learning. Attractive aspects were represented by positive attitudes (ie success and victory). They also mentioned confirmation, verification, and self-reflection ("finding oneself in texts"). These expressions point to self-confidence and cognitive balance. The common denominator of the positive aspects in relevance judgments is the quality of information. One participant said that quality of information followed the "quality of mind".

Study participants agreed that relevance judgment is mainly an intellectual activity. They therefore sought for its positive side in perspective of its successful completion --- finding, discovery, confirmation, verification. The agreeable part of relevance judgment was perceived as a positive value that can be achieved by quality information. Participants also appreciated relevance judgment as being able to select something

45

interesting. They liked moments of surprise and understanding which led to successful problem solving.

5.3   Metaphoric designations of relevance

In order to discover implicit characteristics of relevance, we asked participants to create a metaphor or simile that would express their idea of relevance. They could use whatever expression they found appropriate. Resulting conceptual maps showed that participants used two relations for metaphor creation: *similarity and interaction*. Semantically related concepts expressed similarity as *linking, process, thing (object), abstract value, and person.* The occurrences of expressions related to interaction were based on *contrast*, and *quantity and quality of information.*

The analysis showed that the most frequent metaphors expressed the idea of relevance as *linking, (putting through, fitting in*). They also mentioned that information structures mirror people. For example, "seeking oneself", "mirror", "connection of myself with something else", "wheelwork", "puzzle". An image of seeking was often used. For example, "seeking oneself", "seeking similarity", and "seeking a partner --- intuition". Another repeated image was based on the achievement of a goal. This includes meta-phors such as "seeing the light at the end of the tunnel"and "hitting the target".

Participants also created an objective image of relevance by patterns of ordering or the organisation of information. These patterns bear meanings and can be recognised. For example, "a position in a system" was mentioned. The subjective part of relevance was represented by expressions of intuition, chance, implicit knowledge and skills (eg "target"). In this context the image of *satisfaction* was mentioned: "a good book", "a good drink". Participants also connected relevance with especially valuable, credited information that can be used in practice (eg "inquiry, investigation"). This confirmed relevance as the construction of meaningful information. Other participants expressed attributes of direction and spatial links of information structures (eg "bridge", "foot-print").

The metaphor of "stage for winners" in connection with relevance judgment refers to the need to expand intellectual effort in order to overcome certain obstacles. The basis is "cleaning away" the nonessential, as expressed by the expression "diamond in a stack of coal". Another attribute of relevance was mentioned as creative potential coming from the human mind (eg "something of discovery", "challenge"). The majority of participants used metaphors with a positive aspect, emphasising the goal-oriented activity of a subject. Relevance in action was confirmed by images such as seeking, never-ending run, inquiry, and a road. *Novelty, cleaning and future prospects* have been identified as not often explicitly recognised attributes of relevance.

46

The analyzed metaphors indicate that participants intuitively applied a special language for relevance statements. Metaphoric expressions contain a great deal of implicit information and confirm mostly spontaneous experience in relevance judgments.

5.4 Criteria applied to relevance judgment

Study participants confirmed that they used criteria conditioned by contexts. Criteria usage is activated by stages of research exploration. Contexts were conceptualised mainly as tasks, and the topics of the thesis. Participants also often confirmed the use of citations. Natural contextual factors such as time (timeliness) and space/place (institutions, publishers) were also noted. Basic criteria used for assessing relevance from professional texts were grouped into the following categories: *author, topic, activities and contexts*. Participants also confirmed that they created and used the criteria intuitively. This finding corresponds with similar studies that proved the existence of informal metadata schemas (Koh & Kerne 2006: 303), (Schamber.& Eisenberg & Nilan 1990: 755-770), (Barry 1994: 149-159). The intuitive, informal and incremental character of the process of relevance judgment emerged from the analysis. Intuitive aspects were further explored by questions about supportive criteria used for relevance assessment.

The intuitive criteria were divided into people's subjective states and objective, intuitive factors. Subjective states included categories such as mood, reconstruction of information, inference, cognitive prediction, and first information. Other categories revealed serendipity and unique personal experience. Subjective emotions may have both positive and negative effects, namely lack of time. Participants´ expressions indicate that their relevance judgments depend on their basic mood, attitudes to information, but also on a vague sense of context. Preliminary outlines or schemes of information structuring make the interaction between a person and information easier.

Objective factors of intuitive indicators were determined as document content, document form and authors. In the category of document content, participants mentioned cues such as ways of formulation, and the authors' organisation and interpretation of texts. On the document form they used especially visual and linguistic clues, namely graphics, simple language, style. Participants regarded the author as an important criterion that carries implicit trustworthiness and integrity of information. They also mentioned criteria such as the opinions of certain persons (experts, consultants), one's own experience and knowledge, or empirical data and facts.

Analysis of categories showed that our participants apply multiple criteria in their relevance judgments, something that is confirmed by a number of previous studies (Maglaughlin & Sonnenwald 2002: 327-342). It has also been confirmed that criteria depend on stages of work, as Vakkari and Hakala (2000: 562) discovered. Our findings differ in pointing to intuitive and implicit criteria as the results of situated actions.

47

5.5   Relevance in the electronic environment

The most important finding confirmed that participants invest a great deal of effort into the quest for the origin of the source in the electronic environment. Relevance judgments increase the need for navigation from reference sources to content. Participants mentioned that relevance judgments are supported by advanced technological features of both interfaces and search engines. The added value of the electronic environment is supported by links in contexts, advanced searching and intelligent embedded interfaces. Relevance is scattered and more interactive. New patterns of relevance map the berrypicking principle. Similar contexts were found in a research project linking the scatter of literature and patterns of searching (Vakkari & Talja 2005:.205).

Our doctoral students confirmed that they used electronic sources frequently, but this depends on contexts: topics, disciplines, tasks. In the electronic environment they appreciate accessibility, timeliness, speed as well as technological features such as findability, multimedia form, linking. The majority of participants admitted that they preferred printed sources because of their quality, intellectual analytical depth, stability and predictability. Printed sources were interpreted in a more emotional way, the emphasis being on readability and reliability.

Several students expressed differences between their perception of printed and electronic texts. They agreed that manifold representations make the construction of meanings from electronic texts more complicated. Reduction of stimuli contributes to the problems of relevance in electronic texts. When reflecting on the negative characteristics of electronic sources, students mentioned categories of vagueness and costs or payment. The common denominator of both environments is represented by decision-making frameworks based on verification and reliability of information. Students also discussed time requirements and differences between the quality and quantity of information.

The emerging model of relevance in the networked information environment draws on principles of rich functionality and efficient electronic information processing, which integrates creation, seeking and use. This environment supports relevance by flexibility of navigation and interaction, high-level visualisation, collective information processing and non-linearity. Different grades of contextual support open ways to new relevance patterns in the electronic environment, namely rich verification and trustworthiness of sources.

In line with similar studies (Maglaughlin & Sonnenwald 2002: 327-342), our analyses confirmed that relevance can be supported by visual images and that electronic environment facilitates individual and collaborative information use. Rich user experiences should be based on creativity, joy of use, and discovery.

5.6   Organisation of information

Our study participants admitted that, in the case of relevance judgments, they used spontaneous ways of organising information. Main derived categories related to sorting of information are represented by topics, types of information, and contexts. The mental level (cognition) is interrelated with the physical level (system features, material aids). Participants mentioned folders, tasks, dates. The most important structural relationships used for categorisation include hierarchy, associations and aspects/facets. Both vertical (hierarchy) and horizontal principles (in-links, citations, and related sources) were confirmed.

Three subsets of factors that have an impact on supportive information organisation in relevance judgments were also identified. The first subset consists, in particular, of topics and relationships. The second subset consists of previous knowledge, experience and social networks. The third subset is represented by attributes of information objects (eg identification and content indicators).

The goal of information structuring for relevance judgment is to create contexts and reduce cognitive load. This starts with categorisation (based on clustering) and classification (using hierarchy and linearity). If our analyses are correct, memory plays an important part in all of this. Students confirmed that they combined broader principles of ordering (eg faceted organisation) with more sophisticated principles (eg detailed outline of thesis).

Participants admitted multi-criteria cognitive processing and applied intuitive decision-making frameworks when having to construct meaning out of context. In the electronic environment they use interactions, topical contexts and social communication. Interactive relevance supports the successive emergence of sense in social, organisational and cultural contexts.

The study confirmed that relevance assessment behaviour is focused on contexts. This central finding is in accordance with a lot of information interaction/ retrieval in context studies. Participants interpreted contexts as subjective and objective, intuitive and exploratory. Users of digital libraries need more semantic support of their information needs, tasks, information structures and interpretations. Structural relationships (which should be multiple, combining at least hierarchy and aspects/facets) add to the contexts. Different grades of detailed structuring need to be flexible and available for relevance assessment.
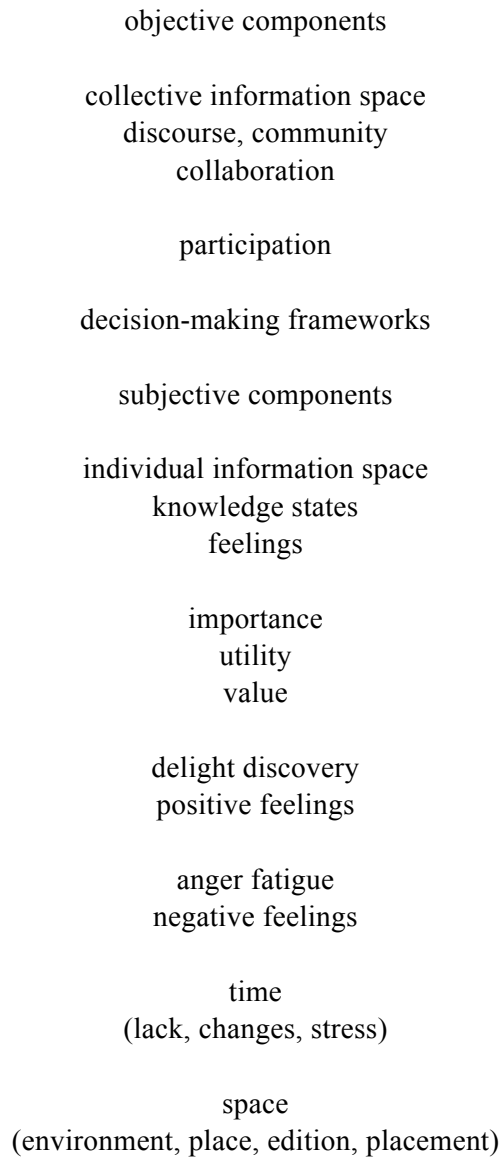
6 Discussion and conclusions

Our study shows that relevance is perceived as value, utility, and importance and the quality of resources plays an important part in all of this. Values point to the benefits of the value sensitive design of systems and services (Friedman & Freier 2005: 368-372). Utility confirms pragmatic contexts of information use that is conceptualised by outcomes of information seeking (Kari 2007). Importance emerges especially from interpretative repertoires investigated by social constructionism (McKenzie 2005: 221-224), (Tuominen & Talja & Savolainen 2005: 328-333). Emotions play a vital role in helping users integrate relevance judgments. We have identified positive feelings of delight and enthusiasm and negative feelings, especially anger. Joy of discovery, self-reflection, and creative moments of surprise are among the most attractive factors of relevance judgment. Intuitive perceptions of relevance implicitly used in metaphors point to novelty, "cleaning away" and planning.

User perceptions and experiences can help develop principles of advanced digital library services. We can see that relevance study points to multiple discovery experiences. It means that, instead of locating information, it is more beneficial to provide the user with features of ranking, relating, recommending and other intellectual activities.

Our study participants confirmed that relevance judgments are based on activities of evaluation, problem solving, decision-making and organisation of information. Analyses disclosed not only explicit, but also implicit parts of relevance judgment. Explicit components are represented by concepts and categories manifested by activities of communications, interactions and behaviour. Implicit components include values, contexts and principles of organisation and structuring.

Further analyses include those aspects of relevance that identify subjective and objective components, two stages of problem solving (initial, concluding), and the factors that influence the process. Perception of relevance and its experience by our study participants are illustrated by a model derived from prior conceptual analyses and categorical models developed by three researchers for single questions (see Figure 1).

objective components

collective information space
discourse, community
collaboration

participation

decision-making frameworks

subjective components

individual information space
knowledge states
feelings

importance
utility
value

delight discovery
positive feelings

anger fatigue
negative feelings

time
(lack, changes, stress)

space
(environment, place, edition, placement)

Figure 1: Conceptual model of perceptions and experience of relevance by study participants

Figure 1 shows the categories derived from our analyses, which are set in the time/space framework (time is indicated by the vertical arrows, space by horizontal arrows). The meanings of the concept of relevance as perceived by participants are outlined in the central nested model (eg the importance of information --- the most frequent meaning). The collective perception in the nested model is surrounded by feelings (positive

51

and negative) as indicated by participants. The boxes of time and space summarise the concepts used most frequently by collective discourse. The left side of the figure divides relevance into the objective and subjective components that emerged from analyses. These result in two important relevance assessment frameworks: participation (networking and collaborative relevance, especially in the electronic environment) and decision-making.

Our findings support the argument about stratified and dynamic contexts of dissertation research which include topic, problem types and accessibility (Vakkari & Pennanen 2001: 217-232),(Chang & Lee 2001: 29-46).This is common with other similar studies(eg Barry [1994:149-159], Byström [2000: 85-101], Chang & Lee [2001: 29-46] and Choo [2007]), in which context emerges as the main driving force of information use. Compared with these and other large-scale studies, our study is based on qualitative methods which have not been applied to the electronic information environment. The findings also suggest that digital libraries research should consider relevance as perception and cognition linked with creativity, relevance as experience integrated by emotions, and relevance as interaction based on communication. The central finding is that the same criteria are used across various types of sources (traditional, electronic) and contexts.

The study of relevance judgments within the environment of digital libraries opens new possibilities for the reconsideration of relevance in contexts of information sharing. Users need support for discovering, decision-making, and participation as part of relevance assessment. For digital library services it is helpful if human cognitive and affective processes are complemented with technological features. There should also be support for the natural behaviour of users manifested by the spontaneous and informal revelation of focus. Non-linearity and uncertainty need tools for backtracking, clues for access to memory, and serendipity. Features which enhance delight and discovery could be beneficial. Supportive values and contexts can help determine an appropriate decision-making framework. In order to help utilisation of information, the transitions between individual and collective information spaces should be facilitated.

References

Anderson, TD. 2005. [online]. Relevance as process: judgements in the context of scholarly research. In *Information Research* January 2005 10 (2), paper 226. Available at: http://InformationR.net/ir/10-2/paper226.html [accessed 30 March 2006]

Anderson T D. 2006. [online]. Uncertainty in action: observing information seeking within the creative processes of scholarly research. In *Information Research*. 2006. 12 (1), paper 283. Available at: http://InformationR.net/ir/12-1/paper283.html [accessed 07 March 2007]

Barry, C L. 1994. User-defined relevance criteria: an exploratory study. *JASIS,* April 45 (3): 149-159.

Bates, MJ. 2005. Berrypicking, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Medford, NJ: Information Today: 58-62**. <<Author: is this the name of the publisher? Please check. J Smith>>**

Blandford, A & Buchanan, G. 2003. [online]. Usability of digital libraries: A source of creative tensions with technical developments. In *TCDL Bulletin*. 2003. Available at: http://www.ieee-tcdl.org/Bulletin/current/blandford/blandford.html [accessed 10 December 2006]

Borgman, CL & Larsen, R. 2003. [online]. In Brief. ECDL 2003 Workshop Report: Digital Library Evaluation --- metrics, testbeds and processes. In *D-Lib magazine* 9 (9) 2003. Available at: http://www.dlib.org/dlib/september03/09inbrief.html#BORGMAN [accessed 07 February 2006]

Bruce, HW. 1994. A cognitive view of the situational dynamism of user-centered relevance estimation. *JASIS* 45 (3): 142-148. **<<Author: month of journal? J Smith>>**

Byström, K. 2000. The effects of task complexity on the relationships between information types acquired and information sources used, in *The new review of information behaviour research: studies of information seeking in context* vol 1 Cambridge: Taylor Graham: 2000: 85-101. **<<Author: if this is a book, we need names of editors please. J Smith>>**

Chang, S-JL & Lee, ZZ. 2001. Conceptualizing context and its relationship to the information behaviour research, in *The new review of information behaviour research: studies of information seeking in context*, vol 2. Cambridge: Taylor Graham: 2001: 29-46. **<<Author: see previous comment. J Smith>>**

Choo, CW. 2007. [online]. Information seeking in organizations: epistemic contexts and contests. In *Information Research*, 2007 12 (2) paper 298. Available at: http://InformationR.net/ir/12-2/paper 298.html [accessed 07 May 2007]

Dervin, B. 2003. [online]. Human studies and user studies: a call for methodological inter-disciplinarity. In *Information Research.* 2003, 9 (1), paper 166. Available at: http://InformationR.net/ir/9-1/paper166.html [accessed 10 November 2005]

Ellis, D. 2005. Ellis´s model of information-seeking behavior, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Medford, NJ: Information Today: 2005: 138-142. **<<Author: is this the name of the publisher? J Smith>>**

Friedman, B & Freier, N. 2005. Value sensitive design, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Meford, NJ: Information Today 2005: 368-372. **<<Author: see previous query. J Smith>>**

Froehlich, TJ. 1994. Relevance reconsidered --- towards an agenda for the 21st century: introduction to special topic issue on relevance research. *JASIS*, April 45 (3): 124-134.

53

Frommholz, I & Fuhr, N. 2006. Evaluation of relevance and knowledge augmentation in discussion search, in *Research and advanced technology for digital libraries.* Proceedings of the 10th ECDL conference. Berlin: Springer: 279-290.

Harter, SP. 1992. Psychological relevance and information science. *JASIS*, October 43 (9): 602-615.

Hjørland, B & Christensen, FS. 2002. Work tasks and socio-cognitive relevance: a specific example. *JASIS* 53 (11): 960-965. **<<Author: month of journal? J Smith>>**

Ingwersen P & Järvelin, K. 2005. *The turn: integration of information seeking and retrieval in context.* Dordrecht: Springer: 448.

Kari, J. 2007. [online]. Conceptualizing the personal outcomes of information. In *Information Research*. 12 (2) paper 292. Available at: http://InformationR.net/ir/12-2/paper292.html [accessed 27 April 2007]

Koh, E & Kerne, A. 2006. "I keep collecting": college students build and utilize collections in spite of breakdowns, in *ECDL 2006: research and advanced technology for digital libraries*, edited by Gonzalo et al. Proceedings of the 10th ECDL conference. Berlin: Springer: 303-314. **<<Author: names and initials of all editors please. J Smith>>**

Kuhltau, CC. 2005a. [online]. Towards collaboration between information seeking research and information retrieval, in *Information research*, 10(2), paper 225. Available at: http://InformationR.net/ir/10-2/paper225.html [accessed 25 April 2006]

Kuhltau, CC. 2005b. Kuhltau´s Information Search Process, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Medford, NJ: Information Today: 2005: 230-234. **<<Author: is this the name of the publisher? J Smith>>**

Limberg, L. 2000. Phenomenography: a relational approach to research on information needs, seeking and use, in *The new review of information behaviour research*, vol 1: 51-68. **<<Author: details of editors please. J Smith>>**

Maglaughlin, KL & Sonnenwald, DH. 2002. User perspective on relevance criteria: a comparison among relevant, partially relevant, and not-relevant judgements. *JASIST* 53 (5): 327-342.

Marchionini, G. 1997. [online]. *Information seeking in electronic environment* . Cambridge: Cambridge University Press: 1997. Available at: http://www.ils.unc.edu/~march/isee_book/web_page.html [accessed 20 April 2004]

McKenzie, PJ. 2005. Interpretative Repertoires, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Medford: ASIST-Information Today: 221-224. **<<Author: is this the name of the publisher? J Smith>>**

Mizzaro, S. 1997. Relevance: the whole history. *JASIS* 48 (9): 810-832. **<<Author: month of journal? J Smith>>**

Saracevic, T. 1996. [online]. Relevance reconsidered, in *Information science: integration in perspectives.* Proc. CoLIS2. 1996: 201-218. Available at:http://www.scils.rutgers.edu/~tefko/ [accessed 20 June 2003]

Saracevic, T. 2004. [online]. Evaluation of digital libraries: an overview. Pres.at DELOS workshop on the evaluation of digital libraries. 2004. Available at: http://dlib.ionio.gr/wp7/ws2004_Saracevic.pdf [accessed 1 March 2007]

Shen, R Vemuri, N. Fan, W & Fox, EA. 2006. What is a successful digital library? in *Research*

and advanced technology for digital libraries. Proceedings of the 10th ECDL conference. Berlin: Springer: 208-219.

Schamber, L, Eisenberg, MB & Nilan, MS. 1990. A re-examination of relevance: toward a dynamic, situational definition. *Inf Process Mgmt* 26 (4): 755-776. **<<Author: please supply full name of journal. J Smith>>**

Spink, AG, Greisdorf, H & Bateman, J. 1998. From highly relevant to not relevant: examining different regions of relevance. In *Inf. Process. Mgmt.* 34 (5): 599-621. **<<Author: see previous query. J Smith>>**

Talja, S. 2005. The domain analytic approach to scholars´ information practices, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Medford: ASIST-Information Today: 123-127. **<<Author: is this the name of the publisher? J Smith>>**

Tuominen, K, Talja, S & Savolainen, R. 2005. Social constructionist viewpoint on information practices, in *Theories of information behavior*, edited by K Fisher, S Erdelez & L McKechnie. Medford: ASIST-Information Today: 328-333. **<<Author: see previous query. J Smith>>**

Vakkari, P & Hakala, N. 2000. Changes in relevance criteria and problem stages in task performance. *Documentation* 56 (5): 540-562.

Vakkari, P & Pennanen, M. 2001. Sources, relevance and contributory information of documents in writing a research proposal: a longitudinal case study, in *The new review of information behaviour research*, vol 2: 217-232. **<<Author: details of editors please. J Smith>>**

Vakkari, P & Talja, S. 2005. The influence of the scatter of literature on the use of electronic resources across disciplines: a case study of FinELib, in *ECDL 2005: research and advanced technology for digital libraries*, edited by A Rauber et al. Proceedings of the 10th ECDL conference. Berlin: Springer: 207-217. **<<Author: names of all editors please. J Smith>>**
--- --- ""

Appendix 1

Questions of semi-structured interviews

1   In your view, what is the relevance of information?

2   Which decisions do your most often take in relevance judgments?

3   Which criteria do you apply in relevance judgments about information from professional texts? *(while reading professional printed documents, articles, papers)*

4   How do you proceed in relevance judgments when dealing with information on the Internet? *(retrieved references in Google, secondary information, list of bibliographic references, full-texts, papers and articles from periodicals)*

    4.1   In judging the relevance of information, do you see any difference between print and electronic? What is this difference?

    4.2   How often do you use *network* electronic resources in your work (study)? Compared with printed resources, you use the network *1* seldom, less frequently --- *2* just as often --- *3* more frequently.

    4.3   Do you prefer using *network* electronic resources to traditional ones?
a. If "yes", why? *(speed, readiness, comfort…?)* b. If "no", why not? *(reliability…?)*

    4.4   Do you publish your scientific papers/articles in electronic form? If you publish or you intend to publish, then:

    a.   Where? *(type of the resource --- electronic journal, personal/institutional/ conference web page…)* b. What percentage *(estimate)* of your overall publishing output does the electronic publishing represent?

    4.5   Do you prefer *network* electronic publishing to traditional publishing? a. If "yes", why? *(speed, readiness, comfort)* b. If "no", why not? *(reliability, "value" of publishing…)*

5   What supporting information do you apply most often in relevance judgments? *(eg important, interesting, known, conforming, useful, peripheral etc)*

6   For you, what three influences are decisive in relevance judgments *(eg requirement, topic, aims, context of use, motivation, knowledge state, mood, situation, time, experience, other)*.

7   If you want to get a basic overview of a certain topic, what types of information and resources do you find relevant? *(eg beginning of a problem solution – databases, terminological sources, textbooks etc)*

8   If you want to derive a solution, analyse and synthesise information, what types of information and sources are relevant for you? *(eg end of the task – professional articles, analyses…)*

9   What other supportive criteria do you apply in relevance judgments regarding a topic? *(eg intuitive clues, unaware traits)*

10   What components or types would you determine within relevance of information? *(eg topical, cognitive, social)*
11   What emotions or moods do you most often experience when practising relevance judgments?
12   For you, what is the most exciting or interesting part of relevance judgments? *(eg discovering new things, experience of new connection, understanding the concept etc)*
13   What metaphor or simile would express your idea of relevance of information?
14   Do you have any comments on the questions asked in this interview?

Thank you

MEASURING LEVELS OF COMPLIANCE WITH LEGAL DEPOSIT LEGISLATION IN SOUTH AFRICA: AN EXPLORATORY STUDY

CECILIA PENZHORN

Department of Information Science, University of Pretoria, 0001 Pretoria, South Africa
cecilia.penzhorn@up.ac.za

RETHA SNYMAN

Department of Information Science, University of Pretoria, 0001 Pretoria, South Africa
retha.snyman@up.ac.za

MARITHA SNYMAN

Department of Information Science, University of Pretoria, 0001 Pretoria, South Africa
maritha.snyman@up.ac.za

Abstract

This article reports on an investigation into the extent to which South African publishers comply with legal deposit. Measuring publishers' level of legal compliance is an important factor in the process of establishing an efficient system of legal deposit in a country, but this study showed such measuring procedures to be lacking. A feasibility study was undertaken, using a variety of methods to check lists of publications, to determine whether they had been received as legal deposit. This investigation proved that measuring publishers' compliance levels is practical and possible in South Africa. The article concludes with a number of recommendations for determining the levels of compliance and with a number of recommended techniques that legal deposit institutions can implement.

1 INTRODUCTION

Legal deposit can be broadly defined as "a government provision which compels producers of all types of publications to deposit a certain number of copies of each

publication in designated or similar institutions" (Jasion 1991:7). It is a system that has been in practice in various countries for centuries and is widely acknowledged as the main instrument for building up and preserving a nation's published heritage (Lariviere 2000).

Legal deposit was introduced in South Africa in 1842, when the British Copyright Act was made applicable to the whole British Empire; according to this Act, a copy of each book published in a British colony was to be deposited with the British Museum (Willemse 1962:78). Later on in the nineteenth century, libraries in South Africa were also designated as deposit libraries. After the establishment of the Republic of South Africa in 1961, legal deposit entered a new dispensation with section 46 of the Copyright Act (Act 63 of 1965), and the provision of copies of publications to the British Museum was officially discontinued (Lor & Geustyn 2001). After various developments over the years, the Legal Deposit of Publications Act, 17 of 1982, was passed. Revision of this Act became imperative with the new developments on the local political front in the early 1990s, and the growing importance, globally and nationally, of audio-visual and electronic media (Lor & Cillie 1997:8). New legislation came into being with the proclamation of the Legal Deposit Act, 54 of 1997, which came into force on 1 July 1998.

Five libraries are designated as repositories for the depositing of books. These are: the National Library of South Africa (Pretoria Campus and Cape Town Campus), Mangaung Library Services (previously the Bloemfontein Public Library), Msunduzi Municipal Library (previously the Natal Society Library Pietermaritzburg), and the Library of Parliament in Cape Town[2]. According to the Legal Deposit Act, the publisher of a document is responsible for depositing those materials subject to legal deposit. Publishers are expected to deposit one copy of a publication with each of the five legal deposit libraries within 14 days of the document first being published (South Africa 1997). In practice, however, depositories do not necessarily receive all the material that has been published nationally.

The revised Legal Deposit Act makes provision for a Legal Deposit Committee to coordinate and promote the implementation of the Act (South Africa 1997). In order to determine the current state of the legal deposit in South Africa, and thus obtain the information needed to carry out its overseeing function, in 2002 the Legal Deposit Committee expressed the need for an inquiry into the many issues and questions surrounding the provision of documents subject to legal deposit.

2   Other places of legal deposit are the National Film, Video and Sound Archives in Pretoria for the depositing of certain categories of non-print documents, as well as at least one place of legal deposit in each of the nine provinces to serve as an Official Publications Depository.

This article reports on the second phase of a larger study, which was undertaken in response to the above request. The first phase of the investigation consisted of a survey aimed at exploring how the attitudes, opinions and behaviours of publishers and legal deposit libraries influence the depositing of material subject to legal deposit. The results of the interviews showed, *inter alia,* a perturbing absence of any procedures in the legal deposit libraries to ascertain publishers' level of compliance with legal deposit. The aim of the second phase of the study, which followed the first, was to obtain information on South African publishers' level of compliance with legal deposit. The second phrase of the study also set out to address the abovementioned shortcoming by identifying and proposing tracing and monitoring mechanisms for future use by the legal deposit institutions.

2 Conducting the research

2.1 Introduction

Although the South African Legal Deposit Act, at the time of its promulgation, was one of the first in the world to make provision for the legal deposit of electronic publications (Bazan 2004; Lariviere 2002), the investigation focused on obtaining figures that reflect the number of books received as legal deposit. This restriction was instituted because the legal deposit system in South Africa was, and still is, focused mainly on print publications, of which books form the greater part. Also, an efficient and successful scheme in the print environment will arguably serve as a good foundation for implementing a programme of legal deposit in the digital environment. Given the need to prepare for the preservation of electronic publications (something that is going to have to be done in the very near future), an investigation into the legal deposit of books was identified as both a practical research area and an issue of immediate concern.

For this type of exploration (into an area not previously investigated), conducting a feasibility study was considered to be appropriate. The aim of the feasibility study was to formally try out and evaluate the use of selected methods and techniques for obtaining compliance figures, to see how well they perform, and to see whether the results of these "tests" could be applied to situations in the future.

2.2 Sampling

Samples for the feasibility study were selected using a combination of purposive and judgement sampling. Because only a small total of elements were to be included for this project, they had to be appropriate for this particular study and had to provide the information required. The following data elements were included in the research sample:

60

- ISBN (International Standard Book Number) lists obtained from the National
Library.
The National Library of South Africa, Pretoria Campus, is the national centre re-
sponsible for the allocation of ISBNs to South African publishers. Lists of ISBNs
with the 0-620 prefix were used for this study. These are special lists of numbers
for once-off publishers or first-time publishers who have not yet established
themselves in the publishing business and have not yet received their own unique
ISBN identifiers. Using these lists made it possible to obtain legal deposit compli-
ance figures from a broad range of publishers, instead of having to target a small
number of individual publishers who had their own block of numbers. In addition,
both the literature and results from the interviews conducted in the first phase of
the research project show that smaller or private publishers are those most likely
to experience problems in complying with legal deposit.

- Book reviews in South Africa's daily newspapers.
All the major South African newspapers are available on microfilm, which is
archived at the Pretoria Campus of the National Library of South Africa. Two
major daily newspapers, *Beeld* and *Cape Argus,* were selected for the study. Both
newspapers, which contain a weekly book review section, are based in major cit-
ies in the country and have large daily readership figures. This inevitably means
that the reviews provide information on a broad spectrum of recent/new publica-
tions, including fiction and non-fiction, books for adults and children, and books
published in both Afrikaans and English.

- Publications in South African libraries as listed on Sabinet Online's SACat data-
base.
Sabinet Online provides a wide range of products and services to libraries and
other institutions and companies in southern Africa. SACat, one of its products,
is a national database of most of the library stock in Southern Africa (Sabinet
Online 2006). Lists of books from the Gauteng and Environs Library Consortium
(GAELIC) were selected from the SACat database. GAELIC is the largest and
most successful academic library consortium in South Africa (FOTIM 2006), and
represents the collections of all Gauteng's major academic libraries, an area not
covered by either the ISBN lists or the newspaper book reviews.

2.3   Time frame

Publications published in 2001 were used to determine whether they had been received
as legal deposit. 2001 was chosen as publication year rather than a more recent year,
because of the backlog experienced, by the National Library, in the compilation of

the South African National Bibliography (SANB). At the time that this phase of the study began in 2006, the SANB was complete up to 2002, and possibly included some publications from the year 2003. It was therefore reasonable to assume that books published in 2001 and sent as legal deposit would be recorded in the SANB by the end of 2002. Smaller publishers do not always manage to publish their books in the same year of obtaining the ISBN. To allow for this eventuality, the 620 ISBN lists for 2000 were also used.

2.4   Control database

The actual SANB published online by the National Library was not the resource of choice against which publications were checked. Because of its user-friendly search facilities, the Publishing Trends Database (PTD) (which contains all publications recorded in the SANB) was used instead. The PTD is an enumerative bibliography originally created by the Publishing Studies section at the Department of Information Science, University of Pretoria, to address the need for quantitative data on book publishing in South Africa (Galloway 2004:115). As a result of a unique arrangement with the National Library, the Department receives information of publications concurrently with them that are in the process of being recorded in the SANB. The PTD thus consists of all publications received as legal deposit, and is updated regularly.

2.5   Collection and analysis of data

Both the book reviews in the newspapers and the publications from the GAELIC lists were scanned for books with South African imprints. These titles, and the titles from the ISBN lists, were all individually checked against the PTD to find out whether they had been received as legal deposit. The number of items received (ie that appeared on the database) was calculated as a percentage of the total number of items checked within each of the above-mentioned categories of data elements used in the study.

As a further exercise, the publishers whose books appeared in the book reviews and on the SACat lists were divided into small, medium and large publishers, according to the categorisation used by the Publishers' Association of South Africa (PASA)[3]. For the purposes of comparison, percentages of items received as legal deposit were then individually calculated for these groupings.

Finally it was decided to look at the compliance percentages of smaller publishers, especially since they were identified as a possible problem area. Percentages obtained for the smaller publishers from the book reviews in the two newspapers and the SA-

3   PASA codes publishers according to turnover band structure indicating the size of publishers; small, medium and large, on which PASA membership fees are based (Galloway *et al* 2005:10).

62

Cat database were compared with the results obtained from the 620 ISBN lists, which included private and once-off publishers.

3 Results of analysis

An analysis of the data indicated that:
- Medium-sized publishers showed the highest compliance figure (76%).
- Large publishers followed with a very similar average of 73%.
- Smaller (commercial) publishers lagged behind quite significantly with a compliance figure of 51%, with the average for the 620 ISBN lists only reaching an average of 29%.
- When this 620 figure is included in the overall calculation, the average compliance figure for small publishers comes to 40%.

These results are shown in the following table and charts:
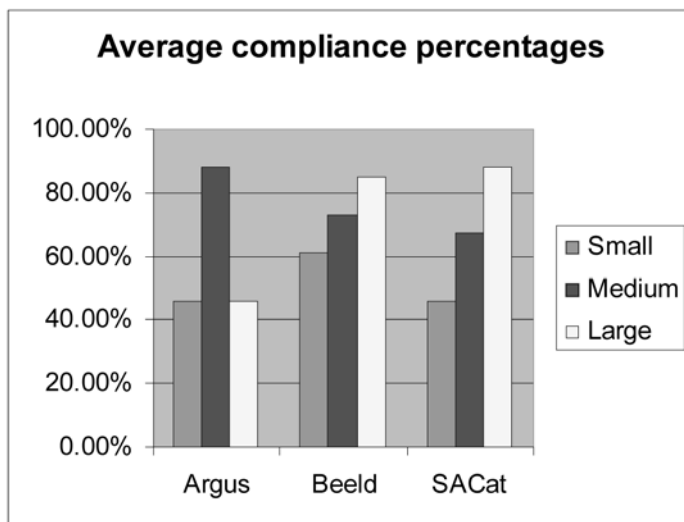
3.1 Average compliance percentages: Table 1

This table shows the percentages of publications received as legal deposit for the year 2001, with the exception of the ISBN lists which reflect percentages for 2000 and 2001 for small, medium and large publishers.

Table 1: Average percentages of publications received as legal deposit

	Small publishers	Medium publishers	Large publishers
Afrikaans newspaper book reviews	61%	73%	85%
English newspaper book reviews	46%	88%	46%
Sabinet SACat database	46%	67%	88%
Average compliance % 620s excluded	51%	76%	73%
620 ISBN lists	2000=30% 2001=28% Average 29%		
Average compliance % 620s included	40%		

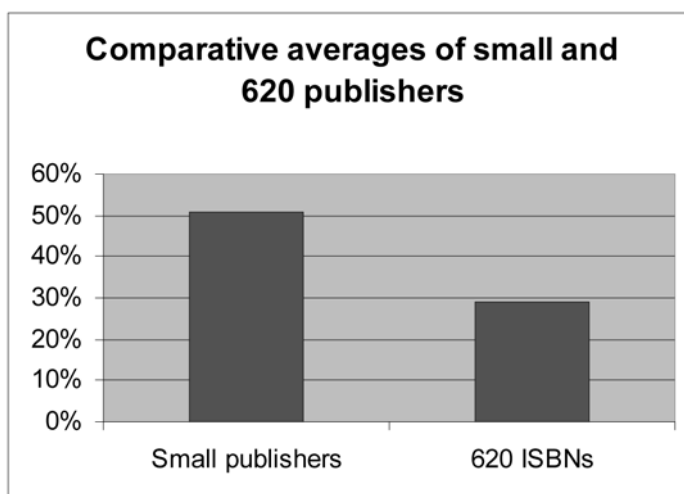## 3.2 Comparative compliance figures: Figure 1

The bar chart in Figure 1 offers a visual comparison of the differences in compliance levels between the small, medium and large publishers as obtained from the studies based on the two daily newspapers and the SACat database. The ISBN lists are excluded because they represent only one category of small publishers.

63

Figure 1: Comparative average percentages of the three categories of publication lists used in the feasibility studies (excluding the 620 ISBNs)

3.3 Comparative compliance levels of smaller publishers: Figure 2

This chart shows the significant differences in compliance percentages between the small (commercial) publishers (listed in the two newspapers and in the SACat database [51%]), and that of the publishers found in the 620 ISBN lists (29%).

Figure 2: Comparison of compliance averages of small commercial publishers and small/ private publishers with 620 ISBNs

64

4   Discussion of results

The results of the feasibility study showed that the legal deposit compliance figures of medium and large publishers are relatively good. The exceptionally low compliance figures of the smaller publishers are a cause for concern and merit serious attention.

The reason for the low percentage of delivery by the large publishers for the English newspaper (46%) in comparison with the other two lists (Afrikaans newspaper 85%, and SACat 88%) is not clear. It does illustrate, however, the need and opportunity for investigation if this type of phenomenon is noticed in a "real life" legal deposit exercise.

There are various possible reasons for the differences in compliance levels between the smaller commercial publishers and the small and/or private publishers with 0-620 ISBNs. Publishers obtaining 0-620 numbers may be less aware of, or less knowledge-able about, their legal deposit obligations. Some publishers might obtain ISBNs with the aim of publishing a specific book, but the book may never actually be published. Also, a number of these small publishers may end up only publishing only two, three or even more years after obtaining an ISBN. The need for continuous follow-up is therefore obvious.

Although the average compliance percentages of 76% and 73% for medium and large publishers respectively may seem satisfactory, they are not as high as comparative figures provided by other countries. Crews, for example, gives some indication of statistics in his 1988 article, in which he reports that the *Bibliotheque Nationale de France* claims to receive 90-100% of all material subject to legal deposit (Crews 1988:564); he also claims that the Swiss National Library, with its voluntary system of legal deposit, claims nearly full compliance from publishers (Crews 1988:567), and that in the United States compliance levels are at 90% (Crews 1988:575).

Although Line (1995:2) contends that the legal deposit of certain types of literature, including commercial literature, is poor in the United Kingdom, other writers claim otherwise. Davies holds that overall cooperation from publishers for the deposit of material has been good (Davies 1998:161) and a BNB (British National Bibliography) MARC survey conducted in the United Kingdom indicated a level of non-compliance of about 10% (Chapman 1997). In a recent e-mail correspondence, John Byford, the legal deposit librarian at the British Library, estimates a receipt of 97% of all material liable for legal deposit (Byford 2005).

In Australia, Whitehead (1995) and Triffett (2006) maintain that compliance with legal deposit is high, in the range of 90-95%. In the Netherlands, a comprehensive study that was done to determine the coverage of legal deposit of material including books, dis-

sertations, periodicals and grey literature (Voorbij & Douwma 1997) showed, in spite of their voluntary system of legal deposit that, with the exception of grey literature, there is a high percentage of compliance --- ranging from 88% to 97%.

5 Recommendations

Legal deposit libraries' lack of structured procedures for obtaining statistical compliance data (a fact that came to light during interviews in the initial research phase) is a cause for concern. The feasibility study, although using only a limited number of data elements, was valuable in showing that measuring compliance with legal deposit is, in fact, possible in South Africa. It is therefore recommended that the five legal deposit libraries commence with developing their own coordinated programme of regular checking for compliance as soon as possible, taking into consideration the following:

- Use of ISBN lists
  The National Library in Pretoria is responsible for providing blocks of ISBNs to publishers. When a series of numbers has been used up, publishers are expected to return their duplicate lists of numbers allocated to individual titles to the National Library (for control purposes). The library also keeps records of all the individual numbers with a 0-620 prefix that are allocated directly to first-time or once-off publishers. Interviewees from the National Library, during the first phase of the research project, indicated that neither the duplicate publishers' lists nor the individual 0-620 lists of ISBNs held by the library are used to check the deposit of legal deposit material. This was confirmed in an informal conversation with the head of the ISBN section at the library, which revealed that she had never personally thought of using ISBNs for control purposes, and that the possibility had never been discussed in any conversations of which she was part.

  South Africa has a long and successful history of ISBN use (Lor & Geustyn 2001), and the percentage of books published in South Africa without ISBNs is, in all probability, very small. The feasibility study demonstrated that using ISBNs to identifythe number of books not sent as legal deposit (and, thus, in effect to pinpoint errant publishers) was both practical and viable.

- Use of a national union catalogue/database
  A national database or union catalogue is a valuable bibliographic resource that is used internationally for legal deposit identification and tracing purposes (Beaudiquez 2001; Bourne 1993:100; Lor & Geustyn 2001). The feasibility study showed that meaningful statistics can be obtained through the use of SACat, a

product of Sabinet Online. It is therefore recommended that, apart from the sample chosen for this study, the various other products and databases available through Sabinet Online be included as possible sources for future checking of publications sent as legal deposit.

- Use of other resources in the general book trade
  The National Library, together with the other legal deposit libraries, will need to identify a wide spectrum of other possible sources against which to compare and verify the delivery of publications and identify non-compliant publishers. Current ad hoc practices must be evaluated for practical implementation, procedures established, and staff assigned to do the necessary work.

The following are a few suggestions for monitoring those activities that use resources in the general book trade. This is by no means a comprehensive list, but can be used as a starting point for further action. The libraries may already be performing some of these tasks informally, such as checking magazines and newspapers, but it is recommended that they co-operatively endeavour to put in place formal procedures, and expand the field by developing new areas of implementation.

- The active utilisation of trade literature such as booklists should be investigated. Libraries can, for example, initiate co-operative ventures with bookshops whereby they regularly receive updated lists of all new acquisitions. Larger national stores could be the responsibility of the National Library and the libraries in a specific city could cover the local stock in smaller stores in the various regions and provinces.

- In the same vein, libraries could consult with the South African Booksellers Association (SABA) to investigate the possibility of obtaining any suitable catalogues (or similar products) from them.

- The feasibility study showed that using information contained in the book reviews that appear weekly in many newspapers is a workable proposition. The responsibility for fulfilling this weekly task could easily be divided amongst the libraries on a regional basis.

- Using various sources for the book trade is another option for monitoring legal deposit. Examples of the use of these resources are found in the United Kingdom (UK) (where Whitaker & Sons' BookBank, and BookData's Bookfind, are used to check for UK imprints [Chapman 1997]), and in the Netherlands, where checks

were done against *Boekblad*, a journal for the book trade, and *Brinkman's Cumu-latieve Catalogus*, which gives a comprehensive overview of Dutch publishing (Voorbij & Douwma 1997). In South Africa we have a similar type of resource in BookData/SAPnet, which offers an extensive database of the latest and most up-to-date records of books published and distributed in South Africa. New records are added on a weekly basis and titles can be searched using the special search facility SA BookSearch (BookData/SAPnet 2005).

• Publishers' catalogues are logical sources for use in the libraries' monitoring activities. It is suggested that this valuable source of information be targeted as a means of keeping up-to-date with all categories of newly published material, and that strategies be examined for "coercing" publishers into regularly providing their catalogues to the legal deposit libraries.

• South African National Bibliography (SANB)
The acquisition of published material by way of legal deposit is intended to serve as the essential foundation for recording, in the form of a comprehensive national bibliography, everything that a country produces (Jasion 1991:7; Lariviere 2000; Lor 1995:97). In South Africa, the National Library in Pretoria is responsible for compiling the SANB and any verification of receipt of legal deposit material thus has to be done against the records contained in the SANB.

A major shortcoming regarding any legal deposit-related activities in South Africa is the fact that the SANB is not up to date. Even if formal procedures are put in place for identifying publications and monitoring compliance with legal deposit, as suggested above, this would be of no practical use without an up-to-date national bibliography. It is therefore strongly recommended that attention be directed to making sure that the processing of new legal deposit material received by the National Library in Pretoria is kept up to date and that current retrospective cata-loguing projects are not neglected.

• Role of the National Library
As the institution responsible for compiling the South African National Bibliogra-phy (SANB), the National Library in Pretoria is the obvious choice for taking the major responsibility for controlling and administering a measuring and monitor-ing programme, but this would not exclude the possibility of sharing some of the checking amongst the other four libraries. Individual libraries can, for example, start a programme of doing their own regular spot checks and communicating the information to Pretoria.

68

A core issue that needs to be addressed if any of these recommendations is to be put into practice is the question of the commitment to legal deposit of the part of the management and staff in the five depository institutions. The successful implementation of a legal deposit programme is an ongoing process that depends on an explicit and strategic focus (Lor 2003:148) that must be executed by dedicated and effective personnel and supported by an adequate infrastructure and committed management.

Lack of staff and finances are two factors that could also be holding back many initiatives for effectively implementing such services. In any organisation, financial decisions obviously impact on the activities undertaken (Williams & Johnson 2004:45). In the same way, the scale and nature of activities and services in libraries are determined by the allocation of resources, which include having enough staff and sufficient money to do the work. Government plays an important role here. The Legal Deposit Committee, as the ministerial committee established to oversee legal deposit in the country, has to fulfil its responsibility of influencing parliamentary debates and decisions with a view to obtaining government recognition and support for its activities.

6 Conclusion

The investigation described in this article provided examples of potentially workable methods and techniques that the legal deposit institutions can use to determine legal deposit compliance levels in South Africa. These methods can be applied or adapted and/or other methods added to obtain relevant information and statistics for developing strategies for tracing and enforcing compliance. Results obtained through such checks/measurements can also be useful in monitoring individual publishers' performances and their depositing patterns. Although other types of publications were not discussed, the importance of such publications (eg newspapers and serial publications) should not be disregarded. Gaps in categories of materials can thus also be identified and possible reasons for non-compliance determined.

There is an indisputable need for an efficient functioning legal deposit system in South Africa. In the digital environment the issues surrounding legal deposit are becoming increasingly complex and it is therefore extremely important that an effective and successful scheme for printed material be in place before embarking on a legal deposit of electronic publications (as South African legislation demands). Measuring the level of publishers' compliance with legal deposit legislation is an indispensable factor in the process of establishing an effective and successful legal deposit system.

References

Bazan, CB. 2004. *Visibility of international recommendations for legal deposit of publications in national legislations. Report to the Standing Committee, IFLA National Libraries Section, Buenos Aires.* Available: http://www.ifla.org/VII/s1/pub/legal_deposit>2004-e.pdf (Accessed 30 April 2006.)

Beaudiquez, M. 2001. *Uses and usefulness of national bibliographies: which perspective?* Paper presented at a workshop given during the 67th IFLA Council and General Conference, Boston, August 16-25.

BookData/SAPnet. 2005. *Bookseller Bulletin.* Available: http://www.bookdatasapnet.co.za (Accessed 30 September 2006).

Bourne, R. 1993. National bibliographies: do they have a future? *Alexandria* 5(2):99-110.

Byford, J. 2005. E-mail to C Penzhorn, 25 July 2005.

Chapman, A. 1997. 1994 revisited: a year in the life of the BNBMARC currency survey. *International Cataloguing and Bibliographic Control* 26(2):41-46.

Crews, KD. 1988. Legal deposit in four countries: laws and library services. *Law Library Journal* 80:551-571.

Davies, J. 1998. Safe deposit: a UK publishing view. *Alexandria* 10(2):159-166.

FOTIM (Foundation of Tertiary Institutions of the Northern Metropolis). 2006. *About FOTIM.* Available: http://www.fotim.ac.za (Accessed 8 September 2006).

Galloway, F. 2004. Notes on the usefulness of publishing statistics for a broader South African book history. *English Studies in Africa* 47(1):109-118.

Galloway, F *et al*. 2005. PASA snapshot industry survey: 2003 report. Pretoria: University of Pretoria, Department of Information Science, Publishing Studies. (Unpublished report). **<<Author: please supply all authors' names. J Smith>>**

Jasion, JT. 1991. *The international guide to legal deposit.* Aldershot: Ashgate.

Lariviere, J. 2000. *Guidelines for legal deposit legislation.* Paris: UNESCO. Available: http://www.ifla.org/VII/s1/gnl/legaldep1.html (Accessed 18 March 2002).

Line, M. 1995. Back to basics for national libraries? Editorial. *Alexandria* 7(1):1-2.

Lor, PJ. 1995. Legal deposit: some issues in the international scene. *Mousaion* 13(1/2):94-111.

Lor, PJ. 1997. *Guidelines for legislation for national library services.* Available: http://www.unesco.org/webworld/nominations/guidelines1_h.htm (Accessed 20 March 2002.)

Lor, PJ. 2003. What prospects for national libraries in Africa? A South African perspective. *Alexandria* 15(3):141-149.

Lor, PJ & Cillie, RM. 1997. Revision and repeal of the Legal Deposit of Publications Act, 1982. *Meta-info Bulletin* 7(3):3-9.

Lor, PJ & Geustyn, M. 2001. The South African imprint: bibliographic control and access. Paper delivered at the International Conference on Colonial and Post-Colonial Cultures of the Book, Grahamstown, 6-8 August. (Unpublished).

Sabinet Online. 2006. Home page: http://www.sabinet.co.za (Accessed 4 May 2006).

South Africa. 1997. *Legal Deposit Act, 54 of 1997.* Cape Town. *Government Gazette* no. 18424, vol. 389, 14 November.

Triffett, A. 2006. E-mail to C Penzhorn, 23 May 2006.

70

Voorbij, H & Douma, P. 1997. The coverage by national libraries of national imprints: a study in the Netherlands. *Alexandria* 9(2):155-166.

Whitehead, D. 1995. *Multimedia preservation – capturing the rainbow.* Paper presented at the 2nd National Preservation Office Conference, Brisbane, Queensland, 28-30 November. Available:

http://www.nla.gov.au/niac/meetings/npo95dw.html (Accessed 23 Sept 2004.)

Willemse, JW. 1962. Die pligeksemplaar-voorreg met besondere verwysing na Suid-Afrika. MA Dissertation. Pretoria: University of Pretoria.

Williams, K & Johnson, B. 2004. *Introducing management: a development guide*. 2nd edition. Amsterdam: Elsevier.

INFORMATION NEEDS AND INFORMATION-SEEKING BEHAVIOUR OF ENGINEERS: A SYSTEMATIC REVIEW

MADELY DU PREEZ

Department of Information Science
University of South Africa
preezm@unisa.ac.za

ABSTRACT

The assumption is often made that engineers do not use conventional information systems or libraries. The purpose of this article is to provide a systematic review of engineers' information needs and information-seeking behaviour. Several models have been developed to study users' information needs and information-seeking behaviour. The Model of Information Seeking of Professionals, developed by Leckie, Pettigrew and Sylvain (1996) provided the framework for this systematic review. The discussion on engineers' information needs and their information-seeking behaviour was based on the following six components of Leckie's model: work roles, associated tasks, characteristics of information needs and factors influencing information seeking, awareness, sources, and outcomes. This review is an attempt to acquire an understanding of the different tasks in which engineers are involved and how these tasks influence their information needs and information-seeking behaviour. This systematic review confirmed the assumption that engineers have specific information needs and that they do, in fact, prefer interpersonal communications to using text-based sources of information. The text-based sources of information used by engineers are generally trade sources. The conclusion highlights issues that still need to be researched.

KEYWORDS

information seeking, information needs, engineers, user studies, user behaviour, systematic reviews

1 INTRODUCTION

Engineers, and specifically consulting engineers, do not seem to use the conventional information systems available in libraries (Ellis 1997:401; Gralewska-Vickery 1976:281; Leckie et al 1996:165, 166). A possible reason for this is that engineers have very spe-

cific information needs that are not easily satisfied by existing information systems or libraries. The complexity of engineers' work may also contribute to the complexity of their information-seeking behaviour (Taylor 1991:227; Fidel & Green 2004:12).

The World Federation of Engineering Organisations' Committee on Engineering Information (WFEO/CEI) (1979:8) defines engineering information as "that information which is used, needed, or generated by engineers in the practice of their profession". While the major part of such information is scientific or technical in nature, engineers' information needs seem to have changed and now include other kinds of information, such as economic, marketing, commercial, social, managerial and legal information (Leckie et al 1996:165-166).

It has been shown that engineers have various reasons for using information sources. King and Griffiths (1991:365-366), for example, found that engineers generally use information in four ways. These include: ways of keeping informed about new techniques, the methods being employed in their field, and the application of this information to specific work-related activities (eg new products that could be used for specific projects). Engineers, like everybody else, also use information to prepare various communications, such as written reports, plans or proposals, journal articles or interpersonal communication in the form of information presentations, consultation or advice. Lastly, engineers may well need information for their professional development or continuing education.

To deepen our understanding of engineers' information needs and information- seeking behaviour, a systematic review of the literature was completed. The information in this review will be tested in the empirical study of a masters' study into the information needs and information-seeking behaviour of consulting engineers. Previous studies on engineers' information needs and use of information were mainly conducted within specific companies and usually within a research and development (R&D) context. Until now, the only survey into engineers' information needs was Ward's survey (2001) of the principal engineers at Ricardo Consulting Engineers. The current review is an attempt to establish what is already known about engineers' specific information needs and the sources they select or use to satisfy such needs.

The information-seeking model developed by Leckie et al (1996), known as the Model of the Information Seeking of Professionals, will provide the framework for this article. The model studies professionals' information needs and information-seeking behaviour according to their different work roles and associated tasks, individual information needs, the factors influencing information seeking, and finally, the outcomes of the information-seeking process.

The discussion on engineers' information needs and information-seeking behaviour will be introduced by a very brief look at what a systematic literature review is and how the literature search was conducted.

2 METHODOLOGY

Literature searches were conducted to retrieve information on engineers' information needs and information-seeking behaviours. Literature searches were conducted on full-text databases, such as Academic Search Premier, Elsevier's ScienceDirect, Emerald, Eric, JStor, Library, Information Science and Technology Abstracts (LISTA), and the library catalogue and the Internet. Internet sources included specific websites of engineering companies and professional engineering councils and associations. The terms that were used to conduct searches were: *engineers*, *information needs* and *information seeking*. Only those information sources that reported on research regarding engineers' information use and needs were chosen. Other sources cited in the selected sources were, however, followed up. Most of the literature that was retrieved reported on research intended to acquire an understanding of the types of information required by engineers during their professional careers as well as research that was conducted by information scientists to develop information needs and information-seeking research models.

3 LECKIE'S MODEL

Research models are graphic versions of the operational aspects of theory (Andrews 2005: 404). Some of the models that were used in this study include Byström and Järvelin's Task-based Information Seeking Model (1995), Ellis' Behavioural model of Information Seeking (Ellis & Haughan 1997); Information Seeking and Using (ISU) Process Model (Cheuk Wai-YI 1998); and Leckie's General Model of Professionals' Information Seeking (Leckie et al 1996).

The General Model of Information Seeking of Professionals developed by Leckie, Pettigrew and Sylvain (1996) was based on their research into the information needs of professionals. In this research, Leckie et al (1996) studied the way professionals' personal circumstances and tasks influence their information-seeking and information-usage behaviour. Figure 1 is a graphical representation of Leckie's model.

Work Roles

Tasks

74

Characteristic of Information Needs

Sources of Information

Information
is sought

Awareness of Information

Feedback

Feedback

Outcomes

Figure 1: Model of the Information Seeking of Professionals (Leckie et al 1996:180)

This model assumes that the roles and related tasks undertaken by professionals in the course of their daily practice prompt particular information needs which, in turn, lead to an information-seeking process. This means that information seeking by professionals is very much influenced by a number of interacting variables that can influence the outcome of the information-seeking process. Leckie's model consists of six components: work roles, associated tasks, characteristics of information needs and the factors influencing information seeking, awareness of information, information sources, and outcomes (Leckie et al 1996; Case 2002:126-129; Leckie 2005; Case 2007:127-129).

4 ENGINEERS' WORK ROLES AND ASSOCIATED TASKS

Work roles and associated tasks are the first two components indicated in Leckie's model. An understanding of the nature of engineers' work roles and associated tasks should provide some insight into their information needs and information-seeking behaviour. The possible impact time and budget have on engineers' information-seeking behaviour will be explored in this section.

4.1 TASKS NEEDING INFORMATION

Ingwersen and Järvelin (2005:73) define a work task as "abstract, objective sequences of actions". Each work task can be divided into subtasks. In engineering, for example, each engineering task will utilise various technological devices or systems and each

75

of these devices is an entity made up of several interdependent parts or subsystems (Wolek 1969:472). There may furthermore be a gap between the engineer's knowledge about the task or the devices that will be utilised and the perceived requirements of the task (Belkin, Oddy & Brooks 1982:62). This knowledge gap constitutes an information need (Byström & Järvelin 1995:192) and this need, in turn, leads to information seeking and retrieval (Ingwersen & Järvelin 2005:73).

Byström and Järvelin (1995:192) observed that the complexity or difficulty of the task is one of the most essential factors affecting task performance. An individual's task or his/her problematic situation requires some specified hypothetical sets of information that form the information requirements of the task (Ingwersen & Järvelin 2005:72). As the task complexity increases, so does the complexity of the information required by the engineers, while the number of useful information sources decreases (Byström & Järvelin 1995:211; Ingwersen & Järvelin 2005:83). Within the context of task complexity, information seeking can be understood as a process in which engineers' understanding of their tasks or problems, their information needs, the relevance criteria and available information evolve (Järvelin & Ingwersen 2004:1).

Engineers are expected to make informed decisions in a number of situations. The choices they make largely depend on their understanding of the context of the task and on their success in obtaining information about this context (Hertzum & Pejtersen 2000:2). This understanding of the relationship between the complexity of a task, information needs and task performance now provides information-seeking researchers with the opportunity to obtain some sort of insight into information users' information-seeking behaviour (Järvelin & Ingwersen 2004:1).

4.2 WORK SITUATION

Leckie et al (1996:179) believe it is necessary to do an in-depth examination of an individual's work to acquire an understanding of his/her information-seeking processes. An understanding of engineers' role and work situation can facilitate an understanding of their information-seeking processes, and help predict their information needs (Wheeler 2004:2) and the process of retrieving the required information (Leckie et al 1996:179).

Engineers' work situations differ from the work situations of other professional groups, such as scientists, in terms of their professional activity, their attitudes, their orientations, and even their typical family backgrounds (Allen 1988:3). Engineers and scientists communicate about their work in different ways (Allen 1988:3). Krulee and Nadler (in Allen 1988:4) found that science students tended to value education as an end in itself, while engineering and management students value it as a means to an end. Ritti (in Allen 1988:4) found that engineers' goals are largely confined to meeting schedules and

76

developing products that will be successful in the marketplace, and helping companies expand their activities. The empirical work of the engineer does not, therefore, encourage the continual integration of new ideas and achievements with existing practices (Wolek 1969:473). Publishing research results is therefore not a goal that will have an influence on the availability of published engineering or technical information.

The nature of the project in which engineers are involved is determined by the subject discipline in which the engineer was trained (eg civil, electrical, electronic, or mechanical). An engineer's specific information needs will also be determined by his or her specific role in a project, whether as the consultant, the contractor, developer or supplier of a device to be used in the said project (Leckie et al 1996:165). Engineers' work may also be contractual or intended for in-house use (which will make the information resulting from that project propriety); this work is always subject to time constraints, and is often conducted in an atmosphere of confidentiality (Gralewska-Vickery 1976:281; Leckie et al 1996:165).

The degree to which a task has been structured by rules and routine may have an effect on the use of information (Choo et al 2000:19). Research by Kwasitsu (2003:5) on the information-seeking behaviour of design, process and manufacturing engineers supports Leckie et al's (1996) research results on how the work roles of health-care professionals, lawyers and engineers influence their information needs. Some engineers also seem to spend more time seeking information at the start of a new project and less time as the project matures, while others continuously seek information throughout the project (Ellis & Haugan 1979:400; Kwasitsu 2003:5). The specific roles and tasks of engineers thus become the determinants of information needs (Leckie et al 1996:165,167).

4.3   TIME AND BUDGET

Time and budget are factors that can obviously influence engineers' task performance (Fidel & Green 2004:4). Time constraints include the time to finish the project and the amount of time engineers have at their disposal to spend on a specific project. Budget constraints include the amount of money available to complete a task or to proceed with a particular task, the larger goals of the organisation, and customer satisfaction. This could have a considerable impact on engineers' information needs, information-seeking behaviour and selection of information sources (this impact will be discussed in more detail in 4.2.5). Engineers also seem to be quite independent as far as taking responsibility for their own tasks is concerned (Fidel & Green 2004:4).

4.4   CHARACTERISTICS OF INFORMATION NEEDS

The third component in Leckie's model is the characteristics that influence professionals' information needs. Information needs arise out of situations pertaining to a specific

77

task that is associated with one or more of the professional's (in this case, engineer's) work roles. The variables that influence or shape the information needs of such professional include individual demographics (age, profession, specialisation, career stage and geographic location), context (situation-specific needs that are prompted internally or externally), frequency (recurring needs or new needs), predictability (anticipated needs or unexpected needs), importance (degrees of urgency), and task complexity (easily resolved or difficult) (Ellis 1997; Leckie et al 1996:182).

The model designed by Paisley (1968) best explains these findings by Leckie et al (1996). Paisley's model sees individuals (in this article, engineers) as members of many communication systems. The unique characteristics of individual engineers will therefore influence all aspects of their life and may well affect their role in work groups. So also may the knowledge they have gained, partly by communication with other people and partly by work experience (Gralewska-Vickery 1976:256-257).

In her study on communication and the information needs of earth science engineers, Grawleska-Vickery (1976) found that the range of information required by engineers in their work varied with career stage. For example, the work of junior engineers will tend to be repetitive; junior engineers they read as a matter of habit, as they did when they were still at university (Grawleska-Vickery 1976:260). Intermediate engineers become supervisors. Their theoretical knowledge needs to be updated and their fields of specialisation are clearly defined. They are frequently called upon to make decisions, which they have to take even if not all factual data are available (Grawleska-Vickery 1976:261). It is at this stage that engineers communicate with a wide spectrum of other engineers; they also tend to read more and attend conferences. Senior engineers have less opportunity for direct use of technical skills but do have a greater need for administrative and business skills (Grawleska-Vickery 1976:261).

Leckie et al (1996:183) point out that an unforeseen or unexpected information need could be relatively unimportant and engineers do not need to find a solution to the problem immediately. However, an unexpected information need could be extremely important and extremely urgent. The level of complexity, combined with the degree of importance and urgency, and the question of whether the information need is anticipated or unexpected, may well influence information-seeking activity.

5 FACTORS AFFECTING INFORMATION SEEKING

The characteristics of engineers' information needs are further influenced by factors such as sources of engineering information and the engineers' awareness of information. These factors are the fourth and fifth components in Leckie's model that need to be explored.

78

5.1   SOURCES OF INFORMATION

Engineering information is not necessarily available in a verbal or textual form (Hertzum & Pejterson 2000:7). Engineers obtain much of the information they require by analysing and decoding physically encoded information, such as the devices they use in a project, or by direct personal contact with other engineers. The availability of physically encoded information can greatly influence how engineers seek for information and the type of information source they will choose to satisfy their information needs (Herzum & Pejterson 2000:7).

5.1.1   Types of information sources

According to Gralewska-Vickery (1976:267), the information sources used by engineers can be classified in various ways: technical or non-technical, project- or profession-oriented, public or private, printed or generated on site, and of enduring or ephemeral value. Engineers' primary source of engineering information is based largely on their personal knowledge or on information contained in books or catalogues (which they keep in their offices) (Shuchman 1981:34).

Vakkari (1999:368-369) found that studies differentiate between external and internal sources, human and documentary sources, or formal (librarians, libraries and other information services) and informal sources of information (peers as sources of information). He distinguishes between internal or personal information sources (eg personal documents, files and memorandums, or colleagues within the engineering firm), and external or impersonal information sources (information sources that are not available within the engineering firm). External or impersonal information sources may include information sources that are acquired from personal subscriptions to journals or personally owned textbooks. These external or impersonal information sources do not really fit into this classification and could be identified with internal sources, provided this division of information sources refers to proximity and the immediate availability of sources, rather than information sources produced by or within the engineering company (Vakkari 1999:369).

5.1.2   Internal information sources

Hertzum's (2002:2) study reports a general agreement in research findings on internal communication (ie personal communications with colleagues working for the same organisation) and costs when it comes to engineers' selection and retrieval of information sources. Research findings on the use of internal information sources have found that:

- Internal communication of any kind is more prevalent in engineering work than communication with external information sources (ie personal contacts or text-based information sources that contain relevant information) (Shuchman 1981:171; King & Griffiths 1991:10; Ellis & Haugan 1997:401; Hertzum & Pejterson 2000:2; Hertzum 2002:2; Case 2002:237; Case 2007:25).
- Engineers tend to rely on their own information and on their colleagues' knowledge rather than on the library and other internal sources (Hertzum & Pejterson 2000:2; Hertzum 2002:2; Case 2002:237; Case 2007:255).
- The use of internal communication could reflect areas of specialised expertise (Anderson, Glassman, McAfee & Pinelli 2001:13).
- Development tasks depend on internal information (Gerstenfeld & Berger 1980:171).
- Perceived expertise may be a better predictor of an engineers' decision to use a particular information source rather than simply a prediction of how the engineer seeks information from a single source (Morrison & Vancouver 2000:4).
- "Word of mouth" is the single, most highly rated source of information (Gerstenfeld & Berger 1980:167; Holland & Powell 1995:10; Case 2002:237; Case 2007:257).
- Core resources are often available in the engineer's own personal collection, or come through colleagues, article reviewing, and other socially mediated channels (Sandstrom 1999:19).

Most engineers favour their own "personal knowledge" of engineering, and "personal experimentation" as sources of information (Case 2002:236; Kranich 2005; Weiss 2005). Bates (2002:8) claims that actively searching for information is a relatively rare act in many people's lives simply because they get so much information in the daily course of events. These findings appear to indicate that favouring "personal knowledge" and "personal experimentation" is a general trend among various user groups.

Literature reports on various case studies that convey information about engineers' use of internal information in particular settings. One of these studies was Hertzum and Pejterson (2000:4)'s study on how engineers at Novo Nordisk archive and utilise their documents for future research. Another example was Ward's survey (2001:171) of the 27 principal engineers at Ricardo Consulting Engineers. These engineers identified their personal memory, personal files, personal books, departmental files, books and databases, other records of previous work (WFEO/CEI 1979:10), the library and other engineers as all being important sources of information. Other information sources identified by these engineers include information acquired from clients, manufacturers and suppliers, manuals and brochures, that is, trade literature (WFEO/CEI 1979:10), conferences, training courses and leisure contacts. Kranich (2005) and Weiss (2005),

80

in their research into engineers working for ABB (a leading company in power and automation technologies), confirmed these research findings.

5.1.3 External information sources

Engineers cannot rely solely on internal information sources when completing information tasks. They also need to use sources that are not available within their company or organisation. These sources are known as *external information sources*. External information sources include personal contacts or text-based sources that can be retrieved from conventional information systems or libraries. Some of the studies reported in the literature indicating engineers' preference for internal sources compared with external sources include studies by Schuchman (198:27-28), Schon (1983:172) , Taylor (1991:237), Frischmuth and Allen (in Taylor 1991:237), Hertzum and Pejterson's (2000:5) and Ward (2001:171-172).

Ward (2001:171-172) found that engineers tend to value informal contacts (ie unofficial contact or contact that is not work-related) with colleagues. These informal contacts complement the use of formal information sources (eg published, text-based information sources, conferences, meetings or formal discussions). A report by Shuchman (1981:27-28) indicated that engineers regard their technical assignments as problems to be worked out at the "bench", since engineering problems require original solutions that are not discussed in the literature. Taylor (1991:237) noted that solutions to engineering problems are dictated by the specifications that may need to be altered, depending on the properties of material, design and time constraints.

Nor did Ward's survey (2001:172) uncover any consistent patterns of library use. He found that library use was adventitious, and driven by immediate, practical needs. The most popular reason for using the library was to consult the database. Major complaints related to the unfriendliness of the way the book stock was arranged, and to the absence of articles (out on loan).

It seems as if engineers generally prefer to use internal rather than external sources of information. There also seems to be a tendency to give preference to personal contacts or interpersonal communication rather than using text-based information sources. The reasons for engineers' preference for personal contacts or interpersonal communication will be explored in the next paragraph.

5.1.4 Interpersonal communication

The research conducted by Anderson, Glassman, McAfee and Pinelli (2002:13-14) found that the most used information source was personal contacts or networks within the organisation. Jain and Triandis (1990:29) explain this trend in that many new ideas

are obtained by talking to people who do similar work within the organisation. A personal observation is that consulting engineers also have discussions with possible suppliers or contractors. For example, they may discuss a specific product to determine whether it would fit into the design of a new project or to find out which new products have become available that can be used in the design or planning of a new project.

Jain and Triandis (1990:30) and Anderson et al (2002:14) suggest that this trend could lead to a problem, since "oral communication may [be] dysfunctional if the actors who exchange communication do not share a common language …" Language, concepts and values tend to be unique to engineering projects and the constant use of this local or project specific language could make communication beyond the project boundaries difficult and prone to misunderstandings (Tushman 1982:357; Jain & Triandis 1990:30).

Holland, Pinelli, Barclay and Kennedy (1991:319) refer to studies that indicate that aerospace engineers devote more time on average to the communication of technical information than to any other scientific or technical activity. These authors also report that there is a strong relationship between the communication of technical information and technical performance at both individual and group levels. This implies that the communication of data, information and knowledge is central to the success of the aerospace innovation process.

Zipperer (1993) identified several reasons for exhibit-design engineers' strong preference for obtaining information from their colleagues:
• Engineers often seek feedback on their ideas or designs, either from trusted experts or as impetus for creative discourse. Katz and Tushman (1979:158) and Anderson et al (2002:14) explain this behaviour when they note that oral communication permits rapid feedback, decoding and synthesis of complex information; this type of communication fits well into settings where most research ideas are yet unformed and difficult to articulate.
• A colleague's memory is often the only access point to filed documents, other than manually searching documents.
• Engineers' close working relationships with their colleagues enables them to choose the person to approach in a given situation based on informal distinctions   such as whether the person is helpful, slow, or inefficient.

Since many engineers work on products and/or items of a proprietary nature (Taylor 1991:237), one can add security to this list.

Fidel and Green's study (2004:5) on how engineers working in a particular organisation sought information found that the chief methods of communication with peers and co-workers were e-mail, phone calls and face-to-face communication.

Some engineers also posted documentation on the Web or on a server (in order to communicate with colleagues, contractors or clients), while some saw meetings as a vehicle for communication. Hertzum and Pejterson (2000:4) found that inter-personal communication proved to be the primary way for engineers to become aware of relevant material.

Case (2002:236; 2007:256) notes that the most relevant information is to be obtained from the engineers' clients and colleagues, that is, from those who are involved in the project or task objectives; this also then tends to reinforce the use of this type of information.

Although engineers prefer personal contacts and interpersonal communications to obtain and share information, they invariably need to use text-based information sources as well. These information sources can also be classified as internal or external sources of information.

5.1.5 Text-based information sources

Anderson et al (2002:14) studied how engineers and scientists select a text from a number of written information sources. They found that the primary determinant was the perceived importance or value of the information source. The primary reason given by the engineers for their choice was prior use of the source, source quality, and the complexity and uncertainty of the task. Task complexity and uncertainty seemed to have been less important reasons compared with the relevance of the source (ie to the task).

An information search does not seem to be negatively affected by the accessibility of text-based information sources. Nor are task complexity and uncertainty consistent factors when choosing written information sources. Differing organisational needs and demands could also be a determining factor in whether an engineer will choose a text-based information source (Anderson et al 2002:14).

Different types of publications (literature) fulfil different information needs (King & Griffiths 1991:365). Trade and news journals, for example, provide fairly general, but newly created, information while scholarly or professional journals are more likely to provide specific information. The literature used by engineers varies according to the field of specialisation (Kremer 1980:124) and tends to be related to their trade. Reports, catalogues, handbooks and trade journals tend to be used, rather than research publications (Case 2002:236; 2007:255; WFEO/CEI 1979:10). Kremer's (1980: 124) study found that engineers consider books and manuals their best sources of information, followed by standards and specifications. Once again, these rankings of information channels varied according to the field of specialisation.

83

Engineering standards are design principles aimed at ensuring an agreed level of quality or attainment and consist of information on features and qualities of products, technology of production, treatment methods, research and measurement methods (WFEO/CEI 1979:9). Standards, such as those developed by the Institute of Electrical and Electronics Engineers (IEEE), the International Standards Organisation (ISO), and the International Telecommunications Union (ITU), prompt the development of new products. These standards are frequently used by research and development engineers working for companies such as ABB (a leading company in power and automation technologies that enable utility and industry customers to improve performance while lowering environmental impact) (Kranich 2005).

Patents are also a type of trade literature and a valuable information source, because they provide new solutions that can be used for practical purposes (WFEO/CEI 1979:9).

Mueller, Sorini and Grossman (2006:2) found that books are critical to engineers working for QUALCOMM (an engineering company which develops technologies and solutions for the wireless communication industry). QUALCOMM librarians re-ported that their engineering clients are comfortable with, and use electronic resources, but have a strong preference for accessing what can be considered more traditional library resources, such as books and other physical items (eg access control systems or computers).

Shuchman (1981:44) discusses the importance of hobby magazines (such as *Popular Science*, *Byte* and *Ham Radio*) as sources of engineering information. This author then (1981:144) suggests that using hobby magazines could reflect a trend towards quick reads and easily understood technical journalism. However, some engineers Shuchman knew (1981:44) claimed that the material in peer-reviewed technical journals could be two or three years old, whereas the hobby magazines were apt to be more current. Case (2002:236; 2007:256) found that journal literature invariably was neither specific nor timely enough for the practical matters at hand, but more useful for monitoring the environment shared with competitors. Other research studies that report similar findings are those by Ward (2001) and Garvey and Griffith (1967). Ward's study (2001:172) praises books for their reliability, but damns them for being out of date. Garvey and Griffith (1967:1013) indicate that formal information channels (eg written information sources) contain information that is relatively "old" as compared with the information disseminated through informal sources (eg personal contacts with a product developer).

The WFEO/CEI (1979:9-10) compares the information value of research reports and visits to other institutions. Research reports consist of information on completed research. Visits to other institutions provide engineers with technical and production information. They can then use this information in their own institution.

Mueller et al (2006:3) observed a heavy reliance on web searching in order to obtain information. They noticed a preponderance of basic searching skills and a certain reliance on only one, or maybe two, search engines. QUALCOMM engineers are often not aware of web resources such as IEEE Xplore, ACM Digital Library, patent and standards tools or speciality search engines such as Google Scholar, CiteSeer or Scirus. QUALCOMM engineers who are aware of these resources often use them in a rudimentary manner, overlooking options such as advanced searching, leveraging index terms, and alerts. QUALCOMM engineers also tend to stick with only one or two paid access resources and are reluctant to explore others they have little experience of (Mueller et al 2006:3).

As indicated in this discussion, there are various types of information sources available to engineers. These sources can be text-based or personal contacts and may be either internal or external to an organisation. Engineers also seem to prefer internal sources of information as well as personal contacts. The reasons why engineers select specific information sources will be discussed in the following paragraphs.

5.1.6 Selection of information sources

Engineers "access information through various channels" (eg colleagues, clients, contractors or suppliers and phone catalogues) and "from various sources" (eg colleagues' reference books and internal memoranda) (Byström & Järvelin 1995:193). From an engineer's point of view, a source contains (or is expected to contain) relevant information, whereas a channel guides (or is expected to guide) the engineer to pertinent sources (Byström & Järvelin 1995:193). Nor is there an absolute difference between channels and sources; a channel may turn into a source and vice versa (Byström & Järvelin 1995:193).

A study conducted by Anderson et al (2001:13) investigated the information seeking behaviour of United States aerospace scientists and engineers. The findings of this study were as follows:

- The engineers followed a pattern consistent with the principle of least effort; they preferred personal collections and oral communications within the organisation. They also conferred with others outside the organisation, and consulted the literature; they consulted library intermediaries only as a last option. Holland, Pinelli, Barclay & Kennedy (1991:330) confirm these findings. Gerstberger and Allen (1968:278) remarked that the only way to get an engineer to make use of professional literature is to bring the literature to the engineer.
- As task uncertainty increased, the search was widened from oral contacts to literature searches and to consulting with library personnel.
- Task complexity was not a factor in sequential choice of an information source.

- Task complexity and task uncertainty were not major factors when written information sources (books, journal articles or Internet sources) were chosen.
- Accessibility and quality were not factors in the aerospace engineers' use of information sources.
- The aerospace engineers' regarded the relevance of the information source to their work as the primary determinant in using a specific information source.

An extensive study at the University of Manchester, England (Gerstenfeld & Berger 1980:167) found no direct correlation between the use frequency of a source and its contribution to the task. Gerstenfeld and Berger (1980:167) consequently suggest that researchers do not necessarily use the best sources for information, but instead those they are familiar with. Gerstberger and Allen (1968:279) agree. They also found that engineers compensate for their lack of attention to quality when selecting an information channel by treating any information they come across with varying degrees of scepticism. In this process, the focus is on technical quality rather than source reliability.

Engineers obviously cannot select an information source if they are not aware of the existence of the source, or they do not know whether the source is available and if they believe the source to be inaccessible.

5.2   AWARENESS OF INFORMATION

Direct or indirect knowledge of various information sources and the perceptions about the information-seeking process, or about the information retrieved, plays a crucial role in the overall information-seeking process (Leckie et al 1996:184-185). The individual engineer's general awareness about information sources and/or the content of these sources may well determine the path that the engineer will take in seeking information. Some important variables identified by Leckie et al (1996:185) are familiarity and previous success (results obtained from strategy or source), trustworthiness (how reliable or helpful), packaging (convenience, usefulness), timeliness (found when needed), cost (relative cost-effectiveness), quality (level of detail, accuracy) and accessibility (relative ease of access). Some of these issues will be discussed in the following paragraphs.

5.2.1 Availability of information

Unlike scientists, whose goal is to publish their research findings, engineers' efforts are directed at producing some sort of material "goods" (eg some physical change or devices) (Allen 1988:8). This information is not in a verbal or textual form unless it is accompanied by some technical documentation. When technological documentation does exist, it is often useful only when the author is available to explain and supplement its content (Allen 1988:10; Taylor 1991:235).

86

Pinelli (1991:20) argues that engineers need information to solve an immediate problem or to make a decision, and their primary output is not information for further research, but to produce a product or service. The information engineers need should therefore be accurate, up to date, reliable and original (Leckie et al 1996:165). There is therefore a strong relationship between the communication of technical information and technical performance at both individual and group levels (Holland et al 1991:319).

There are a few barriers in the production of engineering information. One such barrier can be found in engineers' "legacy to posterity", which is encoded in physical structures (Allen 1988:6). Engineers' primary output is therefore a product or service rather than the publication of technical information or research reports (Pinelli in Wheeler 2001:2). Some intellectual and social effort is required to present technical information in a way that triggers other engineers' attention and gets them constructively involved (Hertzum & Pejterson 2000:6). This implies that the only technical information available in a text-based form is information that someone has had the time (and felt need) to write (Hertzum & Pejterson 2000:7). Engineers' employment also provides a barrier to the production of publicly available engineering information. Most engineers are employed by organisations with a mission (profit, national defence or space exploration) (Allen 1988:6). This necessarily demands a degree of identification, and this works in two ways to exclude the technologist from information communication channels outside his or her organisation. Firstly, engineers are limited, by their companies' requirements, to only work on problems that are of interest to their employers and, secondly, they must refrain from early disclosure of their research in order to maintain their companies' advantage over competitors (Allen 1998:6).

These barriers in the production of text-based engineering information have an influence on both the availability and accessibility of engineering information.

5.2.2 Accessibility

Accessibility refers to the ease with which engineers can gain access to or approach information sources, taking into consideration the social, economic and physical costs of this information usage (Culnan 1985:303; Choo et al 2000:12).

Culnan (1985:304) proposes three dimensions of accessibility: gaining physical access to the information sources, translating an information need or request into a language that is understood by the source (interface dimension), and being able to physically retrieve the potentially relevant information (informational dimension). Choo et al (2000: 13) add a psychological fourth dimension to this list: the embarrassment of revealing one's ignorance or need for assistance.

87

Individual engineers have different perceptions of the concept *accessibility* as far as information systems are concerned (Fidel & Green 2004:3). Fidel and Green (2004:5-6) identified twelve factors that could contribute to engineers' perceptions of the concept *accessibility*. These factors ---

- are sources the engineer knows;
- concentrate several types of information in one place;
- can give the right level of detail;
- save time and have the right format;
- are sources with which the engineer feels comfortable;
- are physically nearby;
- are interactive sources that can be searched by means of keywords or codes;
- are available and accessible; and
- are people who are not too busy to help.

The literature gives different reasons why engineers regard information as accessible. Aguillar (1967) and Taylor (1991:228) indicate a dependence on more personal sources, such as memory, friends, relatives, colleagues and peers. The ease of use and engineers' familiarity with an information source determine an engineer's perception of accessibility (Rosenberg 1967:125; Kremer 1980:123; Leckie et al 1996:167). Engineers will also sacrifice quality to minimise the cost of acquiring information (Rosenberg 1967:125; Leckie et al 1996:167). Other reasons given for accessibility are the frequency of use (Gerstberger & Allen 1968:272; Choo et al 2000: 12; Hertzum & Pejterson 2000:3; Kremer 1980:123) and perceptions of technical quality and channel accessibility (Gerstberger & Allen 1968:272; Choo et al 2000:12; Hertzum & Pejterson 2000:3, Kremer 1980:123).

An information source may be available and accessible to individual engineers, but will not be chosen if the engineer does not regard the source as trustworthy.

5.2.3 Trust

Hertzum (2002) studied the importance of trust in software engineers' assessment and choice of information sources. Tseng and Fogg (1999:41-42) and Hertzum (2002:2-3) identified four types of trust (credibility) on which trust in information sources is based:

- first-hand experience or credibility (a tendency to assess people's expertise and trustworthiness based on their personal interaction with people over time);
- reputation or reputed credibility (asking someone for advice based on a colleague's recommendation);
- simple inspection of surface attributes or surface credibility (judging people by the way they dress or the language they use); and

- general assumptions and stereotypes or presumed credibility (trusting friends to tell the truth).

In line with these four types of trust, Van House, Butler and Schiff (1998:41) found that trust is rooted in communities of practice and that the physical distance between people affects their readiness to trust each other. This concept of trust challenges the least-effort principle, by suggesting that engineers' preference for internal sources, such as personal files and colleagues, could just as much be a preference for sources with a known or easily determinable trustworthiness as it is a preference for information that is easily accessible (Hertzum 2002:3).

5.2.4 Packaging

Packaging of information refers to the format (eg electronic documents or hard copies of documents) or particular mediums (eg personal contact, a journal article, a catalogue or an object) of information. Packaging can therefore also be a significant influence, since engineers might need their information in an electronic format, or need to discuss the engineering problem with colleagues to find a solution to the problem (Leckie et al 1996:185).

5.2.5 Timeliness and cost

Mueller et al (2006:2) observed that engineers often have an information-gathering cycle if the need for information changes drastically during the various phases of the project. Information seeking by this group is generally in response to very specific problems or projects. Engineers also need to obtain the information quickly or within an acceptable amount of time (Leckie et al 1996:185), especially when they are seeking answers to immediate problems (Mueller et al 2006:2).

King (1991:369), Hertzum (2002:2) and Sandstrom (in Jacoby 2005:259-260) regard the amount of time engineers are willing to spend acquiring and reading information found in documents as an indicator of the value of the information. The choice to spend time reading is another indicator of the "value" of the information, which is read; this is similar to the way that the price paid for a journal is an indicator of the perceived value of the journal. Effort costs or cost/time may therefore strongly predict the type of information a person seeks when multiple options are available (Hertzum & Pejterson 2000:6; Morrison & Vancouver 2000:4). However, the information and knowledge gained may have substantially greater value than that indicated by the amount of time (and money) expended in acquiring and reading the information. Individuals who regard the costs of information seeking to be high, seek less information than those who regard the costs as low (Ashford 1986:468).

89

Cost/time is the most frequently cited barrier for seeking both oral and written information (Hertzum & Pejterson 2000:6).

5.2.6 Quality

Quality is the last awareness factor that influences engineers' information needs. A research project of how CEOs in the Canadian telecommunications industry scanned their business environments for information about trends and developments (Auster & Choo 1993:13-14) found that the perceived quality of a source (in terms of its reliability and relevance) was a more important predictor of source use than its perceived accessibility. Pinelli (in Leckie et al 1996:186) found that technical quality was one of the strongest determinants of information usage.

6 OUTCOMES OF THE INFORMATION-SEEKING PROCESS

The outcomes of the information seeking process are the last component in Leckie's General Model of Professionals' Information Seeking.

Information can be used in several ways. Taylor (1991:230) and Choo et al (2000:15-16) identified eight categories to describe the way in which people use information: enlightenment, understanding a problem, instrumental, factual, conformational, projective, motivational, personal or political. The results of engineers' information-seeking processes should culminate in the completion of a tangible project, report or service. Some of the common outcomes or results of information seeking identified by Leckie et al (1996:187) in the literature include providing a service (eg advising a client) or a product (eg constructing a building or designing a television camera), completing paperwork (writing the final reports or tender documents), realising operational benefits (eg the installation and commissioning of a security surveillance system), and achieving certain professional development goals.

7 CONCLUSION

This systematic review focused on the information needs and information-seeking behaviour of engineers. The Model of Information Seeking of Professionals, developed by Leckie et al (1996), provided the framework for the review. The six main components of this model are: work roles, associated tasks, characteristics of information needs and the factors influencing information seeking, awareness of information, information sources, and outcomes.

90

According to the literature, engineers' work tasks are abstract, objective sequences of actions that can be subdivided into subtasks. Engineers are also expected to make informed decisions in certain situations. The engineers' work situations determine their information needs and information-seeking processes. Some of the variables that influence or shape engineers' information needs include: individual demographics, context-specific needs, frequency (recurring needs), predictability (anticipated or unexpected needs), importance and task complexity. The different sources engineers use to find solutions to problems may be internal or external sources, interpersonal sources and text-based or written sources. The systematic review indicated that engineers rely on interpersonal contacts and literature from their trade rather than on the information found in technical journals. The review further indicated that the outcomes of this information-seeking process determine engineers' realisation of their operational benefits, the service they provide, or the successful completion of the paperwork for the engineering project in which they are involved.

Much has been written on the information needs and information-seeking behaviours of engineers. Most of the studies are rather dated, but more recent studies regularly cite the old studies, and the information in these earlier studies are just as valid today as they were when the original research was undertaken. The reason for this can probably be found in the nature of engineers' work and the nature of the technical information they need to complete their tasks.

One issue that still needs to be researched in relation to engineers' information-seeking behaviour is the role of modern information communication technologies (ICTs) in providing this user group with the information they need. This role should not only include the place of the Internet in their information-seeking behaviour, but should also address issues such as the use of email, the use of cellular phones, SMSs and multimedia messages (MMSs).

Another topic that could also be researched is whether engineers from different engineering disciplines follow similar strategies when seeking information.

The literature reported in this review clearly indicated that engineers do have specific information needs and that the conventional information systems found in libraries seem unable to meet their needs. Such an assumption, as well as the validity of earlier findings as discussed in this article, however, needs to be empirically tested. This article seeks to offer a framework for such studies, such as the masters' study on which this review was based.

REFERENCES

Aguillar, FJ. 1967. *Scanning the business environment.* New York: MacMillan.

Allen, TJ. 1988. Distinguishing engineers from scientists, in *Managing professionals in innovative organizations: a collection of readings*, edited by R Katz. Cambridge: Ballinger:3-18.

Anderson, CJ, Glassman, M, McAfee, RB & Pinelli, T. 2001. An investigation of factors affecting how engineers and scientists seek information. *Journal of Engineering and Technology Management* 18(2):131-155.

Ashford, SJ. 1986. The role of feedback seeking in individual adaptation: a resource perspective. *Academy of Management Journal* 29: 465-48.

Auster, E & Choo, C W. 1993. Environmental scanning by CEOs in two Canadian industries. *Journal of the American Society for Information Science* 44 (4):194-203. Also available: http://choo.fis.utoronto.ca/FIS/ResPub/JASIS.pdf (Accessed: 25 June 2004).

Bates, MJ. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online* Review 13(5). Also available: http://www.ischool.utexas.edu/~i385df04/readings/Bates-berrypicking.html (Accessed: 5 February 2005).

Bates, MJ. 2002. Toward an integrated model of information seeking and searching. *The New Review of Information Behaviour Research* 3:1-15.

Belkin, NJ, Oddy, RN & Brooks, HM. 1982. ASK for information retrieval: Part 1. Background and theory. *Journal of Documentation* 38(2):61-71.

Bystr□m, K & Järvelin, K. 1995. Task complexity affects information seeking and use. *Information Processing & Management* 31(2):191-213.

Case, DO. 2002. *Looking for information: a survey of research on information seeking needs and behaviour.* Amsterdam: Academic Press.

Case, DO. 2007. *Looking for information: a survey of research on information seeking, needs and behaviour.* 2nd edition. Amsterdam: Elsevier

Cheuk Wai-Yi, B. 1998. Modelling the information seeking and use process in the workplace: employing sense-making approach. *Information Research* 4(2). Available: http://informationr.net/ir/4-2/isic/cheuk.html (Accessed: 27 February 2005).

Choo, CW, Detlor, B & Turnbull, D. 2000. *Web work: information seeking and knowledge work on the World Wide Web.* Dordrecht: Kluwer Academic.

Culnan, MJ. 1985. The dimensions of perceived accessibility to information: implications for the delivery of information systems and services. *Journal of the American Society for Information Science* 36(5):302-308.

Ellis, D & Haugan, M. 1997. Modelling the information seeking patterns of engineers and research scientists in an industrial environment. *Journal of Documentation* 53(4):384-402

Fidel, R & Green, M. 2004. The many faces of accessibility: engineers' perception of information sources. *Information Processing and Management* 40:563-581.

Garvey, WD & Griffith, BC. 1967. Scientific communication as a social system. *Science.* New series, 157(3792):1011-1016.

Gerstberger, PG & Allen, TJ. 1968. Criteria used by research and development engineers in the selection of an information source. *Journal of Applied Psychology* 52(4):272-279.

Gerstenfeld, A & Berger, P. 1980. An analysis of utilization differences for scientific and technical information. *Management Science* 26(2):165-179.

Gralewska-Vickery, A. 1976. Communication and information needs of earth science engineers. *Information Processing and Management* 12:251-282

Hertzum, M. 2002. The importance of trust in software engineers' assessment and choice of information sources. *Information and Organisation* 12(1):1-18.

Hertzum, M & Pejtersen, AM. 2000. The information-seeking practices of engineers: searching for documents as well as for people. *Information Processing & Management* 36(5):761-778.

Holland, MP, Pinelli, TE, Barclay, RO & Kennedy, JM. 1991. Engineers as information processors: a survey of US Aerospace Engineering Faculty and students. *European Journal of Engineering Education* 16(4):317-336.

Holland, MP & Powell, CK. 1995. A longitudinal survey of the information seeking and use habits of some engineers. *College and Research Libraries* 56(1):7-15.

Hurd, JM, Weller, AC & Curtis, KL. 1992. Information seeking behavior of faculty: use of indexes and abstracts by scientists and engineers. *Proceedings of the American Society for Information Science* 29:136-143.

Ingwersen, P & Järvelin, K. 2005. *The turn: integration of information seeking and retrieval in context.* Dordrecht: Springer.

Jacoby, JA. 2005. Optimal foraging, in *Theories of information behaviour,* edited by KE Fisher, S Erdelez, S & McKechnie. Medford: Published for the American Society for Information Science and Technology by Information Today:259-264.

Jain, RK & Triandis, HC. 1990. *Management of research and development organizations: managing the unmanageable*. New York: Wiley.

Järvelin, K & Ingwersen, P. 2004. Information seeking research needs extension towards tasks and technology. *Information Research* 10(1):1-13. Available: http://informationr.net./ir/10-1/paper 212.html (Accessed: 18 August 2006).

Katz, R & Tushman, M. 1979. Communication patterns, project performance, and task characteristics: an empirical evaluation and integration in an R&D setting. *Organizational Behavior and Human Performance* 23:139-162.

King, DW & Griffiths, JM. 1991. Indicators of the use, usefulness and value of scientific and technical information, in *Online Information 91: 15th International Online Information Meeting proceedings, London, UK, Dec 10-12 1991.* New Jersey: Learned Information:361-377.

Kranich, M. 2005. An informal discussion in Luzern, Switzerland on 7 June, 17:00.

Kremer, JM. 1980. Information flow among engineers in a design company. Thesis (DPhil) – Graduate College of the University of Illinois, Urbana-Champaign.

Kwasitsu, L. 2003. Information seeking behavior of design, process, and manufacturing engineers. *Library & Information Science Research* 25(4):459-476.

Leckie, GJ. 2005. General model of the information seeking of professionals, in Fisher, KE, Erdelez, S & McKechnie, EF. *Theories of information behaviour.* Medford: Published for the American Society for Information Science and Technology by Information Today:158-163.

Leckie, GJ, Pettigrew, KE & Sylvain, C. 1996. Modelling the information seeking of professionals: a general model derived from research on engineers, health care professionals and lawyers. *Library Quarterly* 66(2):161-193.

93

Morrison, EW & Vancouver, JB. 2000. Within-person analysis of information seeking: the effects of perceived costs and benefits. *Journal of Management* 26(1):119-137.

Mueller, BK, Sorini, G & Grossman, E. 2006. *Information seeking behaviour of engineers in the corporate environment: implications for information delivery.* Available: http://www.sla.org/documents/conf/toronto/Mueller.doc (Accessed: 11 June 2007).

Paisley, WJ. 1968. Information needs and uses. *Annual Review of Information Science and Technology* 3:1-30.

Pinelli, TE. 1991. The information seeking habits and practices of engineers. in Steinke, ICA (ed). *Information seeking and communicating behavior of scientists and engineers*. Bingham: Haworth. Also published in: *Science and Technology Libraries,* 11(3):205-209.

QUALCOMM. 2007. Available: http://www.qualcomm.com/ (Accessed on 19 January 2007).

Rosenberg, V. 1967. Factors affecting the preferences of industrial personnel for information gathering methods. *Information Storage and Retrieval* 3:119-127.

Sandstrom, PE. 1999. Scholars as subsistence foragers. *Bulletin of the American Society for Information Science* 25(3):17-20.

Schuchman, HL. 1981. *Information transfer in engineering.* Glastonby: Futures Group.

Taylor, RS. 1991. Information use environments, in Dervin, B & Voigt, MJ (eds). *Progress in communication sciences, vol. 10*. Norwood: Ablex.

Tseng, S & Fogg, BJ. 1999. Credibility and computing technology. *Communications of the ACM* 42(5):39-44.

Tushman, ML. 1982. Managing communication networks in R&D laboratories, in Tushman, ML & Moore, WL. *Readings in the management of innovation*. Marshfield: Pitman

Vakkari, P. 1999. Task complexity, problem structure and information actions: integrating studies on information seeking and retrieval. *Information Processing and Management* 35:819-837.

Van House, NA, Butler, MH & Schiff, LR. 1998. Cooperative knowledge work and practices of trust, in *CSCW '98: proceedings :ACM 1998 Conference on Computer Supported Cooperative Work, Seattle, Washington.* New York: Association for Computing Machinery:335-343.

Ward, M. 2001. A survey of engineers in their information world. *Journal of Librarianship and Information Science* 33(4):168-176.

Weiss, R. 2005. Information management at ABB. [Slide presentation]. [Baden: ABB].

Wheeler, A. 2004. *Review of literature exercise: the information seeking behavior of engineers.* Available: http://www.pages.Drexel.edu/~adw29/my_courses/info511/i511_rol.htm (Accessed on 27/02/2005).

World Federation of Engineering Organisations. Committee on Engineering Information (WFEO/CEI). 1979. *Engineer's needs for scientific and technical information.* Prepared by A. David *et al.* Paris: WFEO/CEI.

Wolek, FW. 1969. The engineer: his work and needs for information. *Proceedings of the Annual Meeting of the American Society for Information Science* 6:471-476.

Zipperer, L. 1993. The creative professional and knowledge. *Special Libraries* 84(2):69-82.

HISTORY OF BIBLIOGRAPHIC CONTROL IN SOUTH AFRICA

INA FOURIE

Department of Information Science, University of Pretoria, Lynnwood Road
Pretoria, 0002, South Africa
inafourie@up.ac.za

MARLENE BURGER

Department of Information Science, University of South Africa
PO Box 392, 0003,Unisa
burgem@unisa.ac.za

ABSTRACT

This article covers the history of bibliographic control in South Africa, with the focus on the legislation regulating legal deposits and a national agency for bibliographic control, and the various role players (eg interest groups, bibliographic networks, the national archival service and service providers for bibliographic control tools). A selection of bibliographic control tools, including union catalogues, bibliographies, indexes and databases, as well as advisory groups and research reports, are also discussed.

1 INTRODUCTION

National bibliographic control activities in South Africa have always been aimed at making texts by South African authors publicly available. These texts include books, journal articles, theses, archival documents, non-print media (Lor & Geusteyn 2003) and other indigenous works. These form part of the heritage of the country that needs to be preserved for future generations. This is why they are identified, described and listed in a systematic manner so that we are *aware* of this information (ie recording existence of these in national bibliographic control tools), so that it can be *located for public use* (ie this refers to physical access or holdings), so that we can accommodate *resource sharing* and *document delivery*, and so that the information sources and texts can be *preserved* for future generations. This is very similar to other countries working towards national and Universal Bibliographic Control (UBC). According to Behrens (1991:42): "Effective national bibliographic control in South Africa ultimately con-

tributes to universal bibliographic control (UBC), and the success of both depends on the cooperation of individuals and institutions in the information community."

A variety of bibliographic control tools and supporting services has been developed since South Africa's earliest attempts at national bibliographic control. According to Musiker (2005), Mendolssohn's *South African Bibliography* (1910) is a landmark work for South African bibliographies. The need for uniform cataloguing and a national South African bibliography was especially stressed in the recommendations following the Carnegie Visit in 1929 (Walker 1993:61). Note that Lor and Geusteyn (2005) offer an interesting review of the bibliographic control of books between 1796 and 1996.

For national bibliographic control to be effective, there needs to be:
* a legal deposit regulation
* a regulating, controlling and advisory body
* supporting committees and investigatory bodies
* a wide selection of bibliographic tools covering various facets of the national recorded heritage
* compliance with national and international standards and certain technological developments (eg computerisation, use of the Internet)
* national support from institutions, consortia, professional associations, interest groups and qualified, dedicated individuals
* research programmes and initiatives
* literature reporting on national bibliographic control

Based on a review of the literature, the above will be used to reflect on the history of South African bibliographic control, which forms the bases for current activities. Since it is not possible to offer a comprehensive review, sources for further information will also be highlighted. The following sources are especially useful: Behrens (1991,1997, 2000), Coetzee (1980, 1991), Coetzee and Boon (1988), Kalley, Schoeman and Burger (2005), Musiker (1996, 2000), Musiker and Musiker (1999), and Viljoen and Zaaiman (1998).

2 Legal enforcement of national bibliographic control

There are two main issues at stake in the legal enforcement of national bibliographic control in South Africa, namely:
* Regulation of legal deposits of all documents.
* Establishment of a national bibliographic agency.

2.1 Legal deposit

For effective national bibliographic control (ie any object which is intended to store or convey information in textual, graphic, visual, auditory or intelligible form), responsible bodies must be aware of all such documents created and published in the country (government documents are particularly important). These responsible bodies should also receive a free copy so that documents can be catalogued, incorporated into bibliographic control tools, and preserved for future generations; also, there needs to be an indication of national holdings. In South Africa, legal deposit, in some form or another, dates back to 1842. It is currently regulated by the Legal Deposit Act, 54 of 1997. All publishers in the country must deposit copies of each document with the following agencies:
- National Library of South Africa (NLSA) (formerly the State Library and the South African Library --- as will be discussed in section 2.2)
- City Library Services, Bloemfontein
- Library of Parliament in Cape Town
- Natal Society Library in Pietermaritzburg.

It is also a legal requirement that the National Film, Video and Sound Archives in Pretoria receive certain categories of documents.

Act 54 of 1997 repealed an earlier Act, 17 of 1982, the Legal Deposit of Publications Act. Lor and Guesteyn (2003) explain the reasons for the new Act, and make a point of mentioning the legal deposit of broadcast media and dynamic (online) electronic media, while Van Zyl (1999) reports on audio visual publications.

The implementation of the Act is coordinated and promoted by a Legal Deposit Committee, which consists of the heads of the places of legal deposit, the head of the Government Printing Works, a representative of all provincial official publications depositories, and two representatives of the publishing industry. The Act allows for the creation of Official Publications Deposits (OPDs) to ensure that government publications are available (De Beer 1998; Lor 2003; Tucket 2003; Tyobeka 1997; Van Zyl 1999a, b). More detail about legal deposits can also be found in Cillie (1997), Coetzee (1999), Fraser (1974), Lor (1995, 2000b, 2001), Lor and Cillie (1998), Marais and Lor (2000), *Report of the first meeting of the new Legal Deposit Committee* (1999) and Willemse (1963). Westra (1997) includes an account of a court case based on a refusal to submit a book.

2.2 National bibliographic agency as regulating, controlling and advisory body

The Legal Deposit Act is closely linked to the forming of a national bibliographic agency. The National Library of South Africa Act, 92 of 1998, legally enforced the

National Library of South Africa (NLSA) (http://www.nlsa.org.za) as the bibliographic agency by amalgamating the former two national libraries: the State Library in Pretoria and the South African Library in Cape Town. The NLSA's responsibilities include:

- Receiving legal deposits as stipulated by the Legal Deposit Act, 54 of 1997, and keeping statistics on the production of published documents.
- Creating bibliographic records, compiling catalogues, bibliographies, indexes and bibliographic databases.
- Attempting to build a complete collection of national documents by adding legal deposits to its collection, and cataloguing and listing these with reference to holdings and preserving these documents for current and future use.
- Allocating International Standard Numbers to publishers in South Africa.
- Acting as a centre for interlending (for countries in Southern Africa).
- Coordinating the compilation of a national periodical index, *Index to South African periodicals (ISAP)*. (Discussed in section 4.2.4.)

For information on the role of the former national libraries, Lor and De Beer (1994) report on the State Library. Information can also be found in Cillie (1996), Fraser (1974), Lor (2000a) and Ritchie (2002), as well as publications of the former national libraries, *Quarterly Bulletin of the National Library of South Africa* and *Quarterly bulletin of the South African Library*.

In 1997 the State Library, the predecessor of the NLSA, established the MARC office to provide training courses and to ensure that libraries can obtain advice on MARC21. MARC training was to be co-ordinated with the Sabinet Online (section 4.3.2) Carnegie-funded training and the IGBIS (section 3.1.4) training. BibSA (Bibliographic Standards in SA) was also established (De Klerk 1998).

Although not part of the NLSA, the South African Library for the Blind (BLIBLIB) in Grahamstown provides national library services to blind and print-handicapped readers.

3 Supporting committees, investigations and seminars

Setting national standards and adhering to these, in compliance with international standards, is important in ensuring effective national bibliographic control. Apart from the role fulfilled by the NLSA and its forerunners, the State Library and the South African Library, a number of specialised committees and advisory bodies were formed and ad hoc investigations conducted into South African national bibliographic control. Behrens (1996a, 1996c, 1997) includes a good overview of this process.

3.1 Committees

3.1.1 SAILIS Committee for Bibliographic Control (CBC)

Under the auspices of the South African Institute for Librarianship and Information Science (SAILIS), the Committee for Bibliographic Control (CBC) initially acted as a standing committee to offer guidelines and recommendations at national level. This committee consisted of the directors of the national libraries and local experts. It was responsible for the SAILIS awards for bibliography, projects on the problems concerning bibliographic control, etc. With the abolishment of SAILIS, most functions of the CBC were taken over by the Interim Committee for Bibliographic Organisations (ICBO).

3.1.2 Advisory Committee on the State Library Bibliographic Services (1994-1996)

The Committee first met in February 1994. It tried to determine what the public at large want regarding bibliographic control. Other activities are described in more detail by Behrens (1997, 1996c). The Committee was disbanded in 1996, with the intention that the ICBO would offer further advice.

3.1.3 Interim Committee for Bibliographic Organisations (ICBO)

ICBO was established in 1996 as an independent committee; it consisted of library and information professionals in South Africa. It had no direct powers, but acted in an advisory capacity in overseeing bibliographic activities in the country. Its functions are discussed by Behrens (1996c, 1997), and included the following:
• Communicating bibliographic organisation needs to the South African government, raising awareness of the importance of bibliographic activities, making recommendations on bibliographic organisation to relevant bodies, and lobbying for funds for such activities.
• Monitoring the national bibliographic tools and prioritising bibliographic projects.
• Promoting standardisation in bibliographic activities.
• Coordinating bibliographic activities in South Africa and southern Africa.

Snyman (1997a, b) reported on an investigation initiated by ICBO on whether UNIMARC or USMARC should replace SAMARC, or whether the latter should be retained. The report of the investigation was submitted to ICBO and all delegates attending a seminar on a future MARC format for South Africa that was held in Pretoria, 24-25 April 1997. Seventy-six percent of the delegates voted in favour of USMARC.

3.1.4 Interest Group for Bibliographic Standards (IGBIS) (2001-)

The Interest Group for Bibliographic Standards (IGBIS) (http://www.liasa.org.za/interest_groups/gbis/php) is a national interest group of the Library and Information Association of South Africa (LIASA), formed in 1997. Although launched in 2001, and properly constituted in 2002, IGBIS has only been active since 2003. IGBIS has no powers of enforcement. Its aim is to promote and facilitate bibliographic standards in South Africa. This includes the following: the promotion of standardisation in bibliographic work at national level, exchanging ideas and experiences (among members) about the application of bibliographic standards, addressing issues which influence the application of bibliographic standards, identifying training needs, conducting continuing education for information workers by arranging talks, meetings and workshops, and supporting academics, researchers, students and the community through the provision of information about bibliographic standards. IGBIS publishes a newsletter, *IGBIS-in-Touch,* which is available to the larger library community.

3.2 Investigations and seminars

Only a few of the most important investigations and interventions are mentioned here.

3.2.1 Inter-University Library Committee and the Fokker report

In 1986, the Inter-University Library Committee (IULC) of the Committee of University Principals requested various task groups to be formed, including the IULC Task Group: Bibliographic Tools. Following a report by the Task Group at the SAILIS national conference of 1991, a new Task Group was formed, whose findings appeared in the Fokker report in 1993. This report commented on the serious commitment of library and information services to South Africa's national bibliographic tools, and the need to keep existing bibliographic tools complete and accurate. It suggested the improvement of certain bibliographic tools and a greater role for both the State Library and Sabinet (Behrens 1997). -

3.2.2 Expert Seminar on Bibliographic Control

In response to the Fokker report, the Expert Seminar on Bibliographic Control was held in May 1994. Participants came from various sectors. Four discussion groups contributed a mission and vision for a national bibliographic service, the functions of such a service, the tools that were required, and how such a service should operate. To pursue the proposals, a Continuation Committee was appointed (Behrens 1997).

100

3.2.3 Behrens report

In March 1995, the Committee for Bibliographic Control started a new investigation into the state-of the-art of the national bibliographic services. The findings of this committee appeared in the Behrens Report. Apart from addressing the state-of-the-art of several bibliographic tools, the report commented on the accessibility of the country's information sources, and the various factors influencing national bibliographic services.

3.2.4 Seminar on Bibliographic Standards

A seminar on Bibliographic Standards for the Promotion of Cooperation was held in February 1996. Several articles were presented on bibliographic standards, the role of tertiary education, in-service training, quality improvement, etc.

3.2.5 Seminar on Accessing Information Resources in Southern Africa

In September 1986, delegates from 14 African countries discussed the need and opportunities for collaboration between the Southern African Development Community countries at a seminar on Accessing Information Resources in Southern Africa.

4 Bibliographic tools

In the following section, a number of South Africa's key bibliographic tools are briefly discussed. We distinguish between national union catalogues of monographs and serials, bibliographies and indexes, databases and a selection of other tools. The following discussion relies heavily on the work of Behrens (2000).

4.1 National union catalogues of monographs and serials

4.1.1 *South African catalogue of monographs, 1941-1975.* Pretoria: State Library, 1976.

Two hundred southern African libraries cooperated in the compilation of this union catalogue for which the former State Library took responsibility. The State Library started the catalogue in 1941 in traditional card form. In 1975, the original catalogue cards were filmed and microfiche sets (consisting of over 2 000 microfiche) of the entire catalogue were made available for sale. Alphabetical codes on the entries indicated the holding libraries. After it closed, the catalogue was superseded by another union catalogue which the former State Library started in 1971; the State Library closed in 1995.

4.1.2 *SA joint catalogue of monographs, 1971 --- 1995.* Pretoria: State Library, 1971-1995.

The *South African catalogue of monographs, 1941-1975,* was superseded by the *SA joint catalogue of monographs, 1971-1995,* which was also the responsibility of the former State Library. It formed part of the *South African cooperative library database (SACD)*, which originated with the establishment of Sabinet in 1983.

The *SA joint catalogue of monographs, 1971-1995,* was published on microfiche in two series, including *Title index* and the *Author index.* Holding statements were in number form. The computerisation of the catalogue was discussed by Van As and Watson (1980). The catalogue continued to be published after Sabinet's new online union catalogue (*SACat*) became available, since not all southern African libraries were members of Sabinet and therefore, for interlending purposes, still needed access to the microfiche version. By 1995 it was no longer feasible to publish this catalogue and the microfiche version also ceased to be produced.

4.1.3 *SACat.* Centurion: SABINET Online, 1983 --- (http://www.sabinet.co.za)

SACat is a current union catalogue of monographs and serials held in South African libraries. It was previously known as the SACD. When Sabinet was established in 1983, it started to create a national database by buying an existing database called WLN from the USA. WLN contained about 500 000 bibliographic records for books and serials. Through two conversion projects, the database of two existing union catalogues for southern Africa --- *Periodicals in southern African libraries (PISAL)* and the *SA joint catalogue of monographs, 1971 --- 1995* were matched against the WLN database.

SACat is used for cooperative cataloguing of both monographs and serials. Members of Sabinet Online either add new records to *SACat* or add their holdings to existing records. *SACat* is also used to find holdings for interlending purposes. Libraries can download records from *SACat* and *WordCat* for inclusion in their own catalogues. Through *SACat* South African holdings are added to *WorldCat*.

SACat Plus, which is available via Sabinet Online's MagNet service, is the name given to the combined databases of the *SACat*, the *SANB*, the *UCTD*, *Navtech* and *Bookdata*.

4.1.4 *Periodicals in southern African libraries (PISAL).* Pretoria: State Library, 1961-1998.

Periodicals in southern African libraries (PISAL) is the national catalogue of serials and was previously compiled by the former State Library (with the cooperation of participating libraries and information services). PISAL was first published under the

102

title *Periodicals in South African libraries* in 1961. In 1977, when holdings of libraries from outside the borders of South Africa were included, the title changed to "southern". Although PISAL was originally published in hard copy, it later became available in microfiche format. It ceased publication in 1998 when it was superseded by SACat. The history of PISAL is traced by Kotze (2005).

4.1.5 Union Catalogue of Theses and Dissertations of South African Universities (*UCTD*). Potchefstroom: Potchefstroom University for Higher Education, 1918 --- .

In spite of its title, this is actually only a bibliographic list of completed titles for post-graduate research at South African universities. Honorary doctorates are also included. It is currently also available through the NEXUS database from NISC (http://www.nisc.co.za/databases?id=7). UCTD is updated annually.

4.2 Bibliographies and indexes

There is a number of national, as well as subject and professional bibliographies.

4.2.1 *South African national bibliography* (SANB). Pretoria: National Library of South Africa, 1959 ---

The NLSA is currently responsible for the SANB using legal deposit documents to compile this national bibliography, which includes monographs (books), pamphlets comprising more than five pages, maps, technical reports, some official publications, periodicals, and the South African Bureau of Standards specifications and microforms. Their cataloguers do the cataloguing online on the *WorldCat*. Initially (1959), SANB was the responsibility of the former South African National Library. It covers the eleven official languages of South Africa. The *SANB* hard copy ceased publication in 1999. It is now available in electronic format through *SACat. SANB* entries are also available on the *WorldCat*, and on the NISC *South African studies* database (discussed in section 4.3.3). At some stage, the SANB was based on PRECIS. This is discussed in more detail by Malherbe (1985). Von Beck (1989) offers an evaluation of the SANB.

4.2.2 *Retrospective South African national bibliography for the period 1926-1958*. Pretoria: State Library, 1985. 2 volumes. (RSANB)

The *Retrospective South African national bibliography* (RSANB) lists sources that were published in South Africa and Namibia during the period 1926 --- 1958. The legal deposit collection of the former State Library was used as the basis of the RSANB, with additional bibliographies used to trace sources not found in this collection. The

RSANB consists of two volumes in hard copy. The first volume contains the bibliographic descriptions for the sources entered alphabetically under the author or, where there is no author, under the title. Each entry is allocated a discreet number, and reference is made to this number in the index, which is contained in volume 2. The index includes authors, editors, compilers, translators, illustrators, series and titles in one alphabetical sequence.

Although considered an urgent responsibility, the retrospective conversion of the South African national Bibliography took 20 years. Initially, it was thought that the project could be completed in a year or two. In his book on the subject, Botha (1985) reflects on the reasons why it took much longer.

4.2.3 A South African bibliography to the year 1925, being a revision and continuation of Sidney Mendelssohn's South African Bibliography (1910). London: Mansell, 1979. Volumes 1-4. Cape Town: South African Library, 1991 and 1997. Volumes 5-6. (*SABIB*).

For many decades, Mendelssohn's bibliography was the most comprehensive retrospective national bibliography for South Africa. In 1960, the former South African Library started with revising the bibliography and adding new sources to the sequence. This was published in 1979 in four volumes as *SABIB*. The first four volumes of *SABIB* include sources published before 1909 that were not listed in Mendelssohn's original bibliography and list sources published up to 1925. A supplement, volume 5, was published in 1991. Volume 6 was published in 1997 and consists of detailed subject and title indexes to the work as a whole.

4.2.4 Index to South African Periodicals (ISAP)

The former Library Association of South Africa started the *Index to South African Periodicals* (ISAP) in 1940. It was continued by the then Johannesburg Public Library in 1945. In 1986 the former State Library acquired authority over ISAP in terms of a ministerial decision. ISAP does not reflect all legal deposit for South African periodicals, but comprises a comprehensive selection of (mainly) scholarly journals and technical and subject general periodicals. A limited number of popular magazines useful in schools or adult basic education are included. The NLSA has contracts with specialised institutions and qualified individuals to prepare records for ISAP. Contributing institutions include:
•   Council for Scientific and Industrial Research (CSIR)

- University of South Africa (UNISA) Law Library
- National Institute for Theology and Religion (NITR)
- Human Sciences Research Council (HSRC)

Kotze (2005) explains the operation of ISAP, while Von Beck (1989) offers an evaluation.

ISAP went online in 1992. Entries date back to 1919, and it is regularly updated by the database providers, Sabinet (sections 4.3.2 & 6.1) and NISC (National Inquiry Services Centre) (section 4.3.3). Kotze (2007) offers a very useful explanation on how ISAP improved their standard and kept up with technological development by using NISCBase indexing and management software. Winter (1967) deals with the history of ISAP.

4.2.5 Specialised bibliographies and indexes

Apart from the national bibliographies and indexes, there is also a number of specialised bibliographies. The history of the bibliography has been documented by Musiker (1980, 1996, 2000). Many of these are also discussed by Behrens (2000). The following are a few examples of significant bibliographies:

- *Bibliography of overseas publications about South Africa, including publications by South Africans and translations of South African works published abroad.* Pretoria: State Library, 1973-181. 5 volumes. This bibliography complements the cataloguing produced by the SANB.
- Peters, MA & Tabane, MM. 1982. *Bibliography of the Tswana language: a bibliography of books, periodicals, pamphlets and manuscripts to the year 1980.* Pretoria: State Library. The Setswana title is *Bibliokerafi ya puo ya seTswana*. Introductory material is written in both   English and Setswana.
- Peters, MA & Bothma, CP. 1992. *Bibliography of the Xhosa language to the year 1990.* Xhosa text edited by GT Sirayi. Pretoria: State Library. The Xhosa title is *Ibhibliyografi yolwimi olusisixhosa*.
- Bothma, CP. *Northern Sotho bibliography to the year 1993.* 1994.   Pretoria: State Library. Northern Sotho language title is: *Bibliokrafi ya Sesotho sa Leboa go fihla ka 1993*.
- Mokhonoana, N & Strassner, M. 1999. *Bibliography of the Zulu language to the year 1998.* Pretoria: National Library of South Africa.
- Musiker, R & Musiker, N. 1997. *Guide to South African reference books*. 6th edition. London: Mannsell.
- Musiker, R. 1996. *South African Bibliography: a survey of bibliographies and bibliographical work.* 3rd edition. London: Mannsell. Contains a list of almost 1 000 bibliographies that deals with various aspects of South Africa.

- *Grahamstown Journal.* Index for 1831 --- 1921, compiled at the Cory Library for Historical Research, Rhodes University, Grahamstown.
- Potgieter, PJJS. 1979. *Index to literature on race relations in South Africa 1910-1975.* Boston, Mass: Hall.
- *A list of South African Newspapers, 1800 --- 1982*: *with library holdings*. 1983. Pretoria: State Library. (Reviewed by Musiker, R. 1984. *South African Journal of Library and Information Science*, 52(1):40.)

4.3   Databases

A selection of databases including previously mentioned union catalogues, indexes, bibliographies and other South African bibliographic control tools are available through Sabinet's *Magnet* (http://www.sabinet.co.za) and NISC's *South African Studies* (http://www.nisc.co.za). There is also *NEXUS*, the national research database.

4.3.1   nexus

The *NEXUS* database of research in progress and research completed is maintained by the National Research Foundation (NRF) (Pretoria) (http://stardata.nrf.ac.za/). *NEXUS* includes several databases, namely:

- Current and completed research projects database, with bibliographic descriptions of more than 70 000 research reports, including masters and doctoral theses. It provides all non-English titles with the English equivalent.
- Talk conference database, with detail on national and international forthcoming conferences.
- Human sciences professional associations database.
- Human sciences research networking database, which lists over 3 000 biographical profiles of research experts.
- Research methodology teachers and courses database.
- Periodical submissions and requirements database.

The records of the Union Catalogue of Theses and Dissertations (UCTD) formerly maintained by the Library of the Potchefstroom University are also included.

Access to the various databases is free of charge on the Web except for the *Current and completed research projects database*. Tertiary institutions have free access to this by means of a user ID and password. Fourie and Van der Berg (2003) discuss the problems experienced with searches on the NEXUS database and the efforts that have been made to improve this search engine.

4.3.2 Databases available through *Magnet* (Sabinet)

A number of databases reflecting South Africa's efforts at national bibliographic control are available through Sabinet's *Magnet* (http://www.sabinet.co.za). These include:

- *SA Gazettes*, which includes the full text of South African Government and Provincial Gazettes and Bills, the index to the *Government Gazette*, as well as the weekly *Government Gazette Index* and the parliamentary bills. Also available is Contemporary Gazette, a fortnightly email notification service and website that focuses on new legislation and Bills that will affect the business world. The possibility of indexing the South African *Government Gazette* was first put forward in 1959. Two initial attempts were unsuccessful. These attempts are discussed by Malherbe (1985). In 1978 the former State Library started with the indexing of the South African *Government Gazette* and the first publication appeared in 1979. It was decided that a book-type index was preferable to a PRECIS system. An in-house set of rules based entirely on pre-coordinate indexing was developed with two indexing tools to establish a controlled indexing language, namely a thesaurus and authority file. Malherbe (1985) reports on these indexing principles and practices.

- *SA tenders*, with full-text access to the Government Tender Bulletin, Provincial Tender Bulletins, private tenders which appear in newspapers and journals and news items from newspapers and journals.

- *SA citation*, plus an index of journal articles, research reports, chapters in books and conference proceedings that are available in South African libraries. The extended ISAP database is included and there are links to holdings in South African libraries. Document delivery options are also provided.

- *SA news* is a combination of two databases, namely the SAPA (South Africa Press Association) database of news reports and the *SA media* database, a collection of press cuttings compiled by the University of the Orange Free State. For *SA media*, most cuttings are from newspapers such as the *Cape Times*, *Cape Argus*, *Rapport*, *Sowetan*, and *Star*. Its predecessor, the press-cutting service of the Institute for Contemporary History, is discussed by Swanepoel (1984).

- *SACat plus* is a combination of bibliographic databases which can be used to support cataloguing activities and includes the *SACat*, the *SANB*, the *UCTD*, *Navtech* and *Bookdata*.

- *SA legal*, as its name suggests, focuses on legal information, including the database of the Commission for Conciliation, Mediation and Arbitration (CCMA) and law reports.

There is also the Sabinet Parliamentary Information (a daily customised alerting service), NetLaw (offering immediate access to updated South African principle acts and statutes), SA Legal (offering access to the latest legislation and judgements) and Law Journal Collection (abstracted records, and full-text documents of many law journals published in South Africa in a single, user-friendly service).

107

4.3.3 South African Studies available through NISC SA

In 1995, NISC SA opened in Grahamstown. NISC SA is involved with the bibliographic activities in South Africa and works to improve national and international access to the national heritage through the databases it publishes and makes available on CD-ROM and on the Internet via the Web-based BiblioLIne (http://www.nisc.com).

South African Studies is an aggregation of 17 databases that provides access to over one million records; some with links to full-text. Records include the majority of documents published in and about South Africa. These include theses, dissertations, conference proceedings, academic and popular magazines, reviews, foreign and local journals and current research. *South African Studies* covers a period from 1960 to the present. In some cases, coverage dates back to the 19th century. Fourie and Behrens (1997) report on an evaluation of *South African Studies*. *South African Studies* includes the following:

- *Index to South African Periodicals*

- *South African National Bibliography*

- The National English Literary Museum (NELM) in Grahamstown, which provides over 247 800 records from 1990 to the present (comprehensive since 1990, retrospective to the 19th century) in six databases:
- *Select Index to South African Literature in English, Critical Writings* which include critical articles and books, theses, dissertations, biographies, bibliographies, and book reviews.
- *Select Index to South African Literature in English, Creative Writings* which covers books, short stories, poems, plays, essays, autobiographies, and travel literature.
- *NELM, Main Catalogue*, which consists of a bibliographic series on famous South African authors.
- *Manuscripts,* including the NELM holdings of literary manuscripts, letters, and recorded interviews. Major collections include: Guy Butler, Hjalmar Thesen, Jack Cope ["Contrast"], Dennis Brutus, Eric Pringle, and others.
- *Literary Awards* offering information on contest rules, prizes, winners, and judges for 280 South African contests and non-South-African contests open to South African entrants.

- *A Bibliography of Anglophone Literature and Literary Criticism by Black South Africans*, including listings of all the creative literature and works of literary criticism published by South Africans of colour between the years 1800 and 1990. This list includes the only first-known instance of the publication of any such work.

The National English Literary Museum in Grahamstown holds a copy of each of the items listed.

- *KnipKat*, a database from the Nasionale Afrikaanse Letterkunde Museum en Navorsingsentrum (NALN) (*National Afrikaans Literary Museum and Research Centre*). This is the indexed database of press cuttings from newspapers and magazines relating to Afrikaans language, literature and culture, and includes information on Afrikaans writers. *KnipKat* provides over 61 800 records, computerised since 1994. NALN also provides a bibliography of books, plays, poetry, manuscripts and sheet music relating to Afrikaans culture.

- Witwatersrand University Management Research Reports, a database compiled by the Witwatersrand Library of Management database, contains over 1 700 records on management in South Africa, from 1970 to the present.

- The Centre for Rural Legal Studies Database covers from 1987 (and earlier) to the present and includes 5 300 records. Subject coverage includes issues relating to agriculture and farm workers in South Africa and elsewhere. It concentrates on labour law and relations, working conditions, housing/tenure, socio-economic conditions and issues such as land reform, rural local government and gender.

- South African Legal Abbreviations, a database of abbreviations which is compiled by the Unisa Law Library (with a view to standardising legal abbreviations).

- Political Information and Monitoring Service Database (PIMS), a database compiled by the Institute for Democracy in South Africa (IDASA). This offers comprehensive, plain language summaries and key details to the Bills, Acts and amendments of the South African Parliament.

- African Journals OnLine (AJOL), which covers Africa's research output.

4.4 Other databases of importance

There are also other databases covering the South African national heritage, including the following:
- *ISALLIS (Index to South African Literature on Library and Information Science)* (1974-). In 1974, the Research Centre for Library and Information Science of the Department of Library and Information Science (now Department of Information Science) at Unisa instituted a research project that led to the development of an indexing system for South African literature on library and information science.

Control over the project was shared between the Department and the Unisa Library. The subjects covered were mostly those taught by library schools in the country and therefore used as a guideline for the scope of the index. Since then, the information scene has changed, with new subjects emerging and curricula expanding to accommodate them. ISALLIS thus has a much broader scope now and includes subjects such as knowledge management and metadata. The following document types are included: all South African journals on Library and Information Science, theses, dissertations, monographs, research reports, conference proceedings, motions proposed at conferences, newspaper articles, general periodicals, inaugural lectures, laws, and so on. South African articles published in overseas journals, chapters in books and individual conference papers are also indexed. There are more than 15 000 indexed items in the database. Each term refers to a brief bibliographic record with an accompanying abstract. The index is available on the Unisa Library database network.

- *Waterlit* (discussed by Myburg 1997) has been produced since 1975 by the Council for Scientific and Industrial Research (CSIR). *Waterlit* covers all aspects of water and water-related subjects, such as aquaculture, environmental issues, hydrology, industrial effluents, legislature, limnology, sanitation and wastewater, all of which is of particular interest to South Africa and developing countries. It is updated weekly with copyright held by the Water Research Commission.

- *FishLit*: 1985 --- current, is compiled by NISC SA in association with the South African Institute for Aquatic Biodiversity Library, Rhodes University Library and Marine and Coastal Management Library. It covers all aspects of fish, shellfish, fisheries and aquaculture literature.

- South African literature in the field of nuclear energy and related fields is contributed to the *International Nuclear Information System (INIS)* (http://www.iaea.org/inisnkm/) by *NECSA* (http://www.necsa.co.za/) formerly the Atomic Energy Corporation of South Africa. Wydeman (1989) provides more detail on the earlier years.

- The DISA project (http://disa.nu.ac.za/) focuses on digital imaging projects in Southern Africa to make historical material concerning the Southern African struggle for democracy more universally accessible.

5 Compliance with national and international standards and technological developments

The importance of standardisation is addressed by Snyman (1996) (among others). This includes standardisation concerning the use of the *Anglo American Cataloguing Rules (AACR2)*, *Library of Congress Subject Headings (LCSH)*, *MARC* and *Dewey Decimal Classification (DDC)*.

Although the conversion of library catalogues through computerisation (eg as discussed by Merrett and Vietzen 1985) is closely linked to bibliographic control, we will focus only on the use of the MARC format and not the conversion history as such. In 1967, the South African National Library Advisory Council (NLAC) took responsibility for the development of a MARC format for South Africa. A subcommittee on Information Retrieval was established to look at developments abroad, and to do a survey on computerised services already available/in development in South Africa. The objective of this was to determine whether libraries in South Africa already participated in international cataloguing services by using magnetic tape services such as MARC. In 1972, the Committee on Bibliographic Services was appointed. They had to do a feasibility study into possible subscription to MARC tapes and similar internationally available services. The study, outcomes, etc are discussed by Coetzee (1997). In 1973, the NLAC appointed a MARC Working Group that replaced the previous working group. In March 1980, the first edition of SAMARC was published by the NLAC. In accordance with the objectives set by Ilse Van Niekerk (her role is explained by Coetzee [1979]) SAMARC was also used as communication format by Sabinet. When Van Niekerk left Sabinet, SAMARC was transferred to the SAILIS Committee for Bibliographic Control (CBC). Roos (1994) and Kingwell (1979) also comment on the use of MARC in South Africa.

In 1997 the South African library community changed to the USMARC format (from SAMARC). Coetzee (1997) discusses this decision and its rationale.

As part of an effort to ensure standardisation in bibliographic control subject description, a number of local thesauri were developed; for example, Merrettt's (1993) *Thesaurus of South African socio-political and economic terms from an anti-apartheid perspective*, as well as the community information thesaurus (Roux & Rykheer 1980).

South Africa also has links with international efforts for bibliographic control and standards (eg through the Sabinet Online Standards Committee SABICAT's involvement with OCLC cataloguing). Sabinet signed a national cataloguing agreement with OCLC, Inc in August 1998, offering South African libraries access to OCLC Cataloguing services, OCLC Interlending services and OCLC PromptCat services. Some libraries, such as the University of South Africa (Unisa), have received Enhanced Cataloguing

Status from OCLC. The Enhance Program is designed to allow skilled cataloguers to improve the quality of the OCLC WorldCat database by upgrading WorldCat records, primarily from less-than-full level to full level. There is also contact through the IFLA Division of Bibliographic Control (http://www.ifla.org/VII/d4/dbc.htm).

6 National support from institutions, consortia, professional associations, interest groups and qualified, dedicated individuals

Support from all affected parties is essential for effective national bibliographic control. Apart from the NLSA, note should also be taken of deposit libraries, and libraries such as Unisa, the role of Sabinet and the South African academic consortia, the National Archives of South Africa, professional associations and interest groups (apart from the advisory committees already discussed) and training service providers.

6.1 Sabinet

Sabinet Online (South African Bibliographic and Information Network) (http://www. sabinet.co.za) was founded in 1983 as a not-for-profit organisation. (Its history is well documented by, amongst others, Boshoff (1980) and Boshoff and Bergenson (1982)). It sold its operational activities in 1997 to Sabinet Online. Its administrative head-quarters is in Centurion near Pretoria. Although initially aimed at supporting national bibliographic control, it currently also caters for commercial and business sectors, and provides access to an extensive list of information sources and bibliographic and full-text databases via the Internet. Apart from consulting services, training and Web publishing, it supports the following:
- Cataloguing services (SABICat).
- Information retrieval services (MagNet) (access to a number of databases– explained in more detail in section 4.3.2).
- Interlending support services (ReQuest). Through ReQuest the requesting library can do the interlibrary loan request and specify the method of document delivery required from the holding library or organisation (eg fax, mail or courier).

6.2 Library consortia

The Coalition of South African Library Consortia (COSALC) (http://www.cosalc.ac.za/) was established on 2 July 1999. Its main focus was access to electronic information through the establishment of the national site licensing initiative, (SASLI), as well as the promotion and support of national co-operative initiatives, such as contributing to the SA National Catalogue (SACat) and interlending agreements, training, capacity

112

building, communication, liaison and lobbying. It is governed by a Board of Directors and includes three membership categories, namely consortium membership, public entity membership and general membership resources. There are the following consortia members which are also linked to the academic consortia:

- CHEC (Cape Higher Education Consortium) linked to <u>CALICO</u> (Cape Library Consortium), established in 1992.
- FOTIM (Foundation of Tertiary Institutions of the Northern Metropolis) linked to <u>GAELIC</u> (Gauteng and Environs Library Consortium), established in 1996.
- esATI (Eastern Seaboard Association of Tertiary Institutions) linked to esAL (Eastern Seaboard Association of Libraries), established in1997.
- ECHEA (Eastern Cape Higher Education Association) linked to <u>SEALS</u> (South Eastern Alliance of Library Systems), established in 1998.
- FSHETT (Free State Higher and Further Education and Training Trust) linked to FRELICO (Free State Library and Information Consortium), established 1996/1997.

- SANRIC (South African National Research Information Consortium), http://sanric.co.za/index.htm established as LISSCO (Library and Information Services of Science Councils) in 1999.

The academic consortia are discussed in more detail by Thomas and Fourie (2006).

6.3   National Archives of South Africa

The National Archives of South Africa (http://www.national.archives.gov.za/) is the coordinating body for union catalogues for archival records. It is situated in Pretoria and has additional repositories in Cape Town, Pietermaritzburg, Bloemfontein and Durban. It undertakes four projects which are the equivalent of the union catalogues:

- *NAROS*: National register of oral sources – audiotapes, audio cassettes, video tapes etc containing interviews.
- *NAREM*: National register of manuscripts: manuscripts, letters, diaries etc (ie hand-written or typed documents).
- *NAREF*: National register of photographs.
- *NAROM*: National register of audio-visual material (ie films, gramophone records, audio tapes, audio cassettes, video tapes and other audiovisual material, music pieces and scores).

The National Archives of South Africa has both a manual (*List of archivalia*) and electronic system (National Automated Archival Information Retrieval System (NAAIRS)). Ferrerai (1985) can be consulted for an overview of the computerisation of the information retrieval system of the Government Archives Section, and Fraser (1981) for

113

a discussion of the subject analysis, cataloguing, preservation, storage and special problems experienced with photographs in archival collections.

6.4 Other professional associations and interest groups

Apart from the role of IGBIS as an interest group of LIASA as explained in Section 3.1.1, the following also contributes to bibliographic control:

- ASAIB (Association of Southern African Indexers and Bibliographers) (http: http://www.asaib.org.za/). ASAIB was established in September 1994 by Professor Reuben Musiker. The association is responsible for two directories, namely *Index of subject specialisation* and *Freelance indexers and bibliographers*. Awards for best indexes and bibliographies are made at ASAIB's annual conference.
- SAMA (South African Museum Association): http://www.samaweb.org.za/
- SASA (South African Society of Archivists): http://www.archives.org.za/
- PASA (Publishers' Association of South Africa): http://www.publishsa.co.za/

6.5 Discussion lists

The SabiCat listserve is maintained by Sabinet for messages concerning issues, problems, etc regarding the SabiCat product and other general comments. There is also the Sabimusic Listserv for music cataloguers in South Africa.

6.6 Training providers

Training forms an essential part of preparing skilled and committed individuals to support national bibliographic control. Behrens (1999:53) stresses the need for such individuals: "This responsibility should manifest itself in each individual: personal attention to cooperation, and personal adherence to standardised bibliographic requirements. Such individual commitment will naturally enhance institutional commitment and cooperation, and thereby show improvement on a national and ultimately international level."

Apart from training offered by tertiary academic institutions, a number of other training opportunities are offered by the NLSA, other libraries, Sabinet Ltd, private companies and individuals. Current curricula in South Africa is outside the scope of this article. Earlier discussion on training and education can be found in Döckel (1989) and Coetzee (1991, 1995).

7 Research activities

In the preceding sections, we covered a number of historical investigations into bibliographic control in South Africa. To ensure a high standard of national bibliographic control it is obviously essential that this form of investigation continues. Many of these appear in the form of Master's or Doctoral studies (eg Willemse 1963 and Myburg 1997). However, current research work is not included in this article, since this is more appropriate for a follow-up article that explores current problems and shortcomings.

8 Literature of national bibliographic control

To ensure continuing effective national bibliographic control we need to understand the past and learn from research efforts. Today, South African literature includes theses, dissertations, articles in national and international periodicals, national as well as international conferences and a number of unpublished reports, minutes of meetings, etc. From the literature study on which this article is based, it has become clear, however, that activities concerning bibliographic control over the past 10 – 15 years especially have not been well documented. This leaves room for an empirical survey to establish what is currently happening in South Africa, and how dissemination of information concerning bibliographic control activities in South Africa can be improved.

9 Conclusion

South Africa has a rich history of bibliographic control. Within the scope of this article we could trace only developments, and have left no room for reflection on problems, shortcomings, etc and how these can shape our knowledge base for future decisions. Understanding these, as well as the full spectrum of our literature is, however, important in ensuring that we can meet future challenges. Following a very brief review of the history of bibliographic control in South Africa, the problems and shortcomings noted will be explored in more detail at the IFLA conference, 19-22 August 2006, <<Author: shouldn't this be 2007? ("will be explored") Please check. J Smith>> Durban. The conference paper, along with this article, will form the basis of an empirical survey that will be conducted with practitioners to determine their perceptions and awareness of problems and shortcomings concerning bibliographic control in South Africa.

References

Behrens, SJ. 1991. National bibliographic control in South Africa. *Mousaion* 9(1):42-54.

Behrens, SJ. 1996a. Expanding access to information: some bibliographic control concerns. *South African Journal of Library and Information Science* 64(2):76-80.

Behrens, S. 1996b. ICBO: Interim Committee for Bibliographic Organisation; Follow-up of the Behrens report, National bibliographic services in South Africa. *Meta-info Bulletin* 5(6):16-18.

Behrens, SJ. 1996c. Recommendations for the improvement of national bibliographic services in South Africa. *South African Journal of Library and Information Science* 64(2):81-85.

Behrens, SJ. 1997. Deliberations on national bibliographic control issues in South

Africa: 1986-1996. *Mousaion* 15(1):50-63.

Behrens, SJ. 2000. *Bibliographic control and information sources*. 3rd edition. Pretoria: University of South Africa.

Boshoff, MC. 1980. 'n Suid-Afrikaanse Biblioteeknetwerk (SABINET): vorderingsverslag. (*A South African Bibliographic Network (SABINET)*. *Suid-Afrikaanse Biblioteke* 48(1):58-67.

Boshoff, MC & Bergesen, AJ. 1982. A tentative tariff structure for the South African Bibliographic and Information Network (SABINET). *South African Journal for Librarianship and Information Science* 49(2):200-209.

Botha, J. 1985. Compilation of the RSANB. *South African Journal of Library and Information Science* 53(4):209-214.

Burger, M. 2002. ISALLIS: from Termatrex to Internet. *The Indexer* 23(2):66-67.

Burger, M & Stinnes, H. 1984. Index tot Suid-Afrikaanse Literatuur oor die Biblioteek- en Inligtingkunde (*Index to South African Literature on Library and Information Science*). *South African Journal of Library and Information Science* 52(2):70-73.

Cillie, RM. 1996. Working Group on the National Libraries of South Africa. *Meta-info Bulletin* 5(4):1-3.

Cillie, RM. 1997. Explanation of the Legal Deposit Bill, 1997. *Meta-info Bulletin* 6(3):16-23.

Coetzee, HS. 1990. Rol en waarde van bibliografiese gesagsbeheer in 'n bibliografiese netwerk. (*Role and value of bibliographic authority control in a bibliographic network*). *South African Journal of Library and Information Science* 58(1):40-46.

Coetzee, HS.1991. Bibliografiese ontsluiting in die jare negentig. (*Cataloguing and classification in the nineties.*) *South African Journal of Library and Information Science* 59(4):282-293.

Coetzee, HS. 1995. Opleiding van studente in katalogisering, klassifikasie en die verlening van onderwerpstoegang tot inligting. (*Educating students in cataloging, classification and subject analysis.*) *South African Journal of Library and Information Science* 63(3):148-158.

Coetzee, HS. 1997. Development of SAMARC: South African format for the exchange of machine readable bibliographic descriptions. *South African Journal of Library and Information Science* 65(3):139-150.

Coetzee, HS. 1999. The Dublin Core initiative as a metadata standard for information retrieval on the Internet. *Meta-information Bulletin* 8(4):9-20.

Coetzee, HS & Boon, JA. 1988. Gehaltebeheer in 'n bibliografiese network (*Quality control*

in a bibliographic network). *South African Journal of Library and Information Science* 56(4):299-308.

De Beer, JF. 1998. Workshop to enhance services of the State Library, Pretoria, and of the official publications depositories. *Meta-info Bulletin* 7(3):10-12.

De Klerk, T. 1998. BibSA (Bibliographic Standards in South Africa): a new section of the State Library, Pretoria. *Meta-info Bulletin* 7(4):34.

Döckel, H. 1989. Die veranderende rol van die ontsluiter in 'n bibliografiese network. (*The changing role of the cataloguer in a bibliographic* network). *Mousaion* 7(2):34-43.

Ferrerai, SJ. 1985. Die Staatsargiefdiens se gerekenariseerde inligtingherwinningstelsel. (*The computerised information retrieval system of the Government Archives Section.*) *South African Journal of Library and Information Science* 53(4):203-208.

Fourie, I & Behrens, SJ. 1997. *South African Studies*: a review of NISC's anthology CD-ROM. *South African Journal of Library and Information Science* 65(2):116-123.

Fourie, I & Van der Berg, H. 2003. A story told by Nexus transaction logs: what to make of it. *Mousaion* 21(2):11-40.

Fraser, HM. 1974. The national library, with special reference to its bibliographical function. MA dissertation, Potchefstroom University for Christian Higher Education.

Kalley, JA, Schoeman, E & Burger, M. 2005. *Indexing for Southern Africa: a manual compiled in celebration of ASAIB's first decade: 1994 --- 2004.* Pretoria: Unisa Press.

Kingwill, HF. 1979. A history of the MARC project in South Africa. *South African Libraries* 47(1):5-14.

Kotze, A. 2005. Periodicals indexing, in *Indexing for Southern Africa: a manual compiled in celebration of ASAIB's first decade: 1994 --- 2004*, edited by JA Kalley, E Schoeman & M Burger. Pretoria: Unisa Press: 92-102.

Kotze, A. 2007. Report on ISAP on NISCBase. (Unpublished).

Lor, PJ. 1995. Legal deposit: some issues in the international scene. *Mousaion* 13(1/2):94-111.

Lor, PJ. 2000a. Role of the National Library of South Africa in ensuring access to research results and data collections among organs of state. *Meta-info Bulletin* 9(3):10-15.

Lor, PJ. 2000b. Modernising legal deposit legislation: the South African experience. *Meta-info Bulletin* 9(4):17.

Lor, PJ. 2001. Legal deposit in South Africa and access to published information. *Meta-Info Bulletin* 10(3):1.

Lor, PJ. 2003. Official publications depositories: work in progress. *Mousaion* 21(1): 6-26.

Lor, PJ & Cillie, RM. 1998. Revision and repeal of the Legal Deposit of Publications Act, 1982. *Meta-info Bulletin* 7(3):3-9.

Lor, PJ & De Beer, JF. 1994. The State Library in the service of South African science and scholarship. *South African Journal of Science* 90(8):462-466.

Lor, PJ & Geusteyn, M. 2003. Expanding bibliographic control and accessibility of non-print media. *Quarterly Bulletin of the National Library of South Africa* 57(3):102-104.

Lor, PJ & Geusteyn, M. 2005. From a trickle to a torrent: bibliographic control of books in South Africa, 1796 to 1996. *Mousaion* 23(1):19-38.

Malherbe, PM. 1985. *Government Gazette Index* --- an index of the *Government Gazette* of

the Republic of South Africa. *South African Journal of Library and Information Science* 53(2):95-101.

Merrett, C. 1993. *Thesaurus of South African socio-political and economic terms from an anti-apartheid perspective*. 3rd edition. Pietermaritzburg: University of Natal.

Merret, C & Vietzen, C. 1985. Aspects of card catalogue conversion at the University of Natal Library, Pietermaritzburg. *South African Journal of Library and Information Science* 53(2):89-94.

Musiker, R. 1980. *South African bibliography: a survey of bibliographies and bibliographic work*. 2nd edition. Cape Town: David Philip.

Musiker, R. 1997. *Guide to South African reference books*. 6th edition. London: Cassell.

Musiker, R. 2005. History of indexing in South Africa: a brief overview, in *Indexing for Southern Africa: a manual compiled in celebration of ASAIB's first decade: 1994 --- 2004*, edited by JA Kalley, E Schoeman & M Burger. Pretoria: Unisa Press: 3-6.

Myburgh, F. 1997. An overview of trends in bibliographic database production methods with specific reference to the Waterlit database. MA dissertation, University of Pretoria.

Report on the first meeting of the new Legal Deposit Committee, 10 February 1999. *Meta-info Bulletin* 8(2):10-12.

Ritchie, G. 2002. The culture of collecting: the National Library as a memory institution. *Historia* 47(2):511-530.

Roos, A. 1994. The South African MARC format (SAMARC). *International Cataloguing and Bibliographic Control* 23(2):26-28.

Roux, PJA & Rykheer, JH. 1990. Developing a South African master thesaurus for community information. *South African Journal of Library and Information Science* 58(2):174-179.

Snyman, MMM. 1997a. Future machine-readable cataloguing format for South African libraries. *South African Journal of Library and Information Science* 65(3):151-159.

Snyman, MMM. 1997b. *Future MARC format for South African libraries: final report*. Pretoria: Department of Arts, Culture, Science and Technology.

Snyman, R. 1996. Importance of bibliographic standards, codes and systems for cooperation. *Meta-Information Bulletin* 5(2):5-6.

Swanepoel, JJ. 1984. Die knipseldiens van die Instituut vir Eietydse Geskiedenis: totstandkoming en werksaamhede. (The press-cutting service of the Institute for Contemporary History: establishment and activities.) *South African Journal of Library and Information Science* 52(4):123-127.

Tyobeka, EM. 1997. The rationale for establishing Official Publications Depositories (OPDs) in South Africa. *Meta-info Bulletin* 6(4):8-10.

Van As, ML & Watson, JA. 1980. The computerisation of the Southern African Joint Catalogue of Monographs in a single sequence. *South African Libraries* 47(3):85-96.

Van Zyl, J. 1999a. Report on the meeting with publishers of audio-visual publications regarding legal deposit, held in Pretoria on 29 June 1999. *Meta-info Bulletin* 8(3):20-26.

Van Zyl, J. 1999b. Report on the Workshop on Official Publications Depositories (OPDs), held in Pretoria on 7 July 1999. *Meta-info Bulletin* 8(3):17-19.

Viljoen, AJ & Zaaiman, RB. 1998. Kernverantwoordelikhede van die Suid-Afrikaanse

inligtinggemeenskap teenoor nasionale bibliografiese beheer. (*Core responsibilities of the South African information community towards national bibliographic control*). *South African Journal of Library and Information Science* 56(3):82-88.

Von Beck, M. 1989. South African bibliographic control tools. *Mousaion* 7(1):76-86.

Walker, CM. 1993. From Carnegie to NEPI: ideals and realities. *Mousaion* 11(2):58-83.

Westra, P. 1997. Legal deposit for the preservation of our published heritage: development. *Cape Librarian* 41(4):48-49.

Willemse, J. 1963. The legal deposit prerogative, with special reference to South Africa. MA dissertation, University of Pretoria.

Winter, JS. 1967. The Index to South African Periodicals. *South African Libraries* 34(3):85-89.

Wydeman, LA. 1989. Die rol van Suid-Afrika in die Internasionale Kerninligtingstelsel (INIS). (*The role of South Africa in the International Nuclear Information System (INIS))*. *Mousaion* 7(1):2-12.

119

National Archives and The Effective Management of Public Sector Records In Kenya

Henry N Kemoni

School of Information Sciences, Moi University, Eldoret, Kenya
hkemoni@yahoo.com

Patrick Ngulube

Information Studies Programme, University of KwaZulu-Natal, Pietermaritzburg, South Africa.
ngulubep@ukzn.ac.za

Abstract

The article presents the findings of a study which examined the role of Kenya National Archives and Documentation Service (KNADS) in the management of records in 18 government ministry headquarters in Nairobi, Kenya. Data was collected (in 2005) through questionnaires sent out to 157 registry staff, and through interviews conducted with 10 senior government officers and six archive personnel from Nairobi Records Centre and KNADS headquarters. The research findings revealed that, even though KNADS provided records management advice to government departments, the record creating agencies did not adequately implement this advice.

The study established that KNADS faced various challenges in providing records management advice to government departments and that this may compromise public service delivery and the effective management of government information. The challenges that were identified included: inadequate human and financial resources, lack of support from senior government officers, low priority accorded to records management in government departments, lack of regular follow-ups on recordkeeping practices in departments, and inadequacies in existing records and archives legislation. The study concluded that KNADS had not effectively helped record-creating agencies to properly manage their records (largely owing to the constraints it faced). Among the key recommendations of the study is that, given the limited resources available, there is need to review existing records and archives legislation in order to divide responsibilities for recordkeeping between KNADS and government departments.

Keywords:

Kenya, National Archives, Public records, Records Management, Public Service Delivery, Recordkeeping

1 Introduction and background to the study

In many countries, including Kenya, national archival institutions are mandated by records and archives legislation that stipulate the proper management of records in the public sector. Thurston (1996:187) underscored this point by observing that national archival institutions in Africa had a statutory responsibility for records management in the public sector and thus, any attempt to understand the development of records management in the public sector in Africa needed to focus on the national archives.

National archives are involved in the management of public sector records, since records management is the key to performance in the public sector. For example, effective records management improves service delivery, supports efficient information exchange, facilitates evidence–based policy making, supports administration of data protection principles, supports the effective implementation of Freedom of Information legislation, encourages accountability, and improves decisions and knowledge management across sectors of government (Blake 2005). This partially explains why Ngulube and Tafor (2006:57) claimed that records and archives contain information that is, in effect, the cornerstone of government accountability and good governance.

A literature search indicated that archival institutions worldwide were involved in the management of public sector records. In the United Kingdom, the National Archives (2004) advises government departments and the wider public sector on best practices in records management while in Australia, the National Archives (2004) provides advice to government agencies by developing policies, standards, guidelines, and providing training and advice about modern recordkeeping. In the United States of America, the National Archives and Records Administration (2004) helps to preserve the nation's history by overseeing the management of all federal records. The key mission of the National Archives and Records Service of South Africa (2004) is to foster national identity and protect human rights by promoting efficient, accountable and transparent government through the proper management and care of government records. The National Archives of India (2005) is currently engaged in streamlining the management of public records.

The government of Botswana (2007) points out that the mission of the Botswana National Archives and Records Services Department is to provide efficient and effective economic management of all public records throughout their life cycle (from creation to disposition), and to preserve those public records of archival value for posterity and access purposes. On the other hand, the Kenya National Archives and Documentation Services (2000), in line with its mission, advises public offices on the proper management of records. The records management responsibilities of KNADS are further spelt out in the existing archival legislation. According to section 4 (1) a

of the Public Archives and Documentation Service Act Cap **<<Author: should this be Chap (short for Chapter? Please check. J Smith>>** 19 (of the laws) of Kenya, the Director of KNADS or officers under him/her, have a mandate to examine public records and advise on their care, preservation and custody.

The records management advice provided by national archives to government departments is important, since record-keeping facilitates effective management of recorded information throughout its life cycle, and may lead to both socio-political and economic development. Recordkeeping may also help in the attainment of the renowned United Nations Millennium Development Goals (UN MDGs). According to the United Nations (2005), the eight UN MDGs include: eradicating extreme poverty and hunger, achieving universal primary education, promoting gender equality and empowerment of women, reducing child mortality, improving maternal health, combating HIV/AIDS, malaria and other diseases, ensuring environmental sustainability and developing a global partnership for development. Access to complete and accurate records partly holds the key to the achievement of some of the UN MDGs. Well-managed information is critical to sound decision-making and socio-economic development.

Good recordkeeping practices may also lead to the attainment of the New Partnership for Africa's Development (NEPAD) objectives. NEPAD is a continental strategic development programme initiated by African leaders such as President Thabo Mbeki of South Africa, Olusegon Obasanjo of Nigeria and Abdelasiz Bouteflika of Algeria. It is a collective action to solve Africa's economic problems and recognises the need to meet the UN MDGs and targets adopted by the United Nations in September 2000 (Okumu 2002:227; Ilorah 2004:223-224).

2 Statement of the problem

Public sector record management programmes in Africa are plagued by various problems due to the ineffectiveness of registries and national archival institutions (Mnjama 2005; Ngulube & Tafor 2006). Central to the problem of managing public records is the fact that archival and records legislation is not conducive to records management development; also, the tradition of organisational support for recordkeeping systems is weak. Consequently, many public offices in Africa have experienced record management problems.

In Kenya, for example, reports of lost files, missing and the misfiling of records were described as a common feature in many government departments (Obudho 1999:2). The Office of the President (1999) noted that cases of missing and lost records were a common experience in public offices. That was caused by poor record management

practices in government departments. Terer (2000:1), the then Permanent Secretary in the Office of the Vice-President and Ministry of Home Affairs, pointed out that missing and lost files and documents resulted in delayed service to the citizens, and projected a poor image of the public service. In recognition of the key role that national archives play in the management of government records, Terer (2001) virtually implored citizens to make formal complaints, in writing, to the Director of KNADS whenever services they required were unduly delayed as a result of missing or lost files and documents. The Director of KNADS was required to submit quarterly reports of such cases for further action.

As records management practices continued to deteriorate, the Government of Kenya appointed a Task Force in the year 2003 to investigate the causes of poor record-keeping in government departments. According to a draft document from the Office of the President, Directorate of Personnel Management (2003a:2), some of the Task Force's terms of reference included analyzing the records management situation in government departments, analyzing the role played by registries in the management of records in the public sector, establishing the constraints that influenced the performance of registries; and preparing an action plan on how to address all the constraints identified. However, the terms of reference of the Task Force did not completely put KNADS under the spotlight.

The unsatisfactory state of record-keeping in the Kenyan public sector was also highlighted by Musila Musembi (2004), and the then Director of KNADS during the ESARBICA Executive Board Meeting held in Maputo in 2003 (ESARBICA 2003). The unsatisfactory record management situation resulted in delayed and poor service delivery, frustrations on the part of the public, and also encouraged corruption. What is the cause of Kenya's poor records management practices? Are registries failing to play their role as gatekeepers of the management of government information? Or is it that KNADS is failing to effectively fulfil its mandate?

In view of the prevailing state of recordkeeping in Kenya, the study investigated the extent to which KNADS effectively carried out its mandate in relation to assisting registry staff to properly manage records throughout the records continuum. The study was guided by the following four research questions:
- What type of record management advice does KNADS provide to government departments?
- Do government departments implement the records management advice provided by KNADS?
- What challenges does KNADS face in providing records management advice to government departments?
- How can the identified challenges be addressed?

123

3 Significance of the study

The current study is significant in a number of ways. Records provide evidence of human activities and transactions, underlie the rights of individuals and states, and are fundamental to democracy and good governance (ICA 2004). In view of the role of records management in enhancing democracy and good governance, the study examined the role KNADS played in providing recordkeeping advice to government departments. KNADS' contribution is vital, since record-keeping underpins all aspects of public administration.

Very few studies in records and archives management reported in the literature are based on empirical evidence (Ngulube & Tafor 2006:58). In this regard, the study findings are significant because they are based on empirical research. The findings would be of use to all record and archive management scholars, educators, consultants, researchers and students in Africa undertaking studies related to the role of national archives in the management of public sector records. This study also supplements previous research within the East and Southern Africa Regional Branch of the International Council on Archives (ESARBICA) (of which Kenya is a member) that have explored the role of national archives in managing public sector records. It will thus provide useful comparable data for scholars and researchers alike.

The findings also offer valuable lessons to other archival institutions in the ESARBICA region on the need to effectively manage government information, irrespective of format. This study will also add to the literature and data on record and archive management, and make some contribution towards records management theory, practice, methodology and policy formulation.

4 Short research story

The study utilised a survey research methodology to investigate the role that KNADS played in managing records in government ministries' headquarters in Nairobi, the capital city of Kenya. The study was based on the assumption that the records management practices at the headquarters of government registries were a good indicator of how those ministries managed their records at all levels of government.

Survey research appears to be a popular method among researchers in Library and Information Science (Ngulube 2005:131). In conducting a records management survey, researchers may use existing documentation, direct observation, questionnaires and interviews as data collection methods (Penn, Pennix & Coulson 1994:59; Shepherd & Yeo 2003:33-35). Data was collected through the use of an interview schedule, a questionnaire, and direct observation (which supplemented data obtained from inter-

views and questionnaires). These data collection strategies were also used by Wamu-koya (1996) to determine the impact of records management on administrative reform programmes in Kenya. The data collection methods used in this study were therefore deemed to be appropriate.

The researcher conducted a survey of records management practices in 18 out of 24 government ministries headquarters as constituted before February 2005. No sampling of government ministries was done. The remaining six ministries were not covered in the study because of the lack of research endorsement by the respective accounting officers. The target population of the study consisted of 173 respondents who consisted of 157 registry personnel, 10 senior government officers who were directly involved in the management of records in their various ministries, and 6 archives personnel from Nairobi Records Centre and KNADS headquarters (which is the watchdog of the management of government records in Nairobi). Questionnaires were administered on registry personnel while interviews were conducted with senior ministerial officers and archives personnel. The observation technique was employed to collect data from the registries, and was used to verify data obtained from senior government officers and registry personnel.

The data obtained was both qualitative and quantitative. The quantitative data that were obtained from the questionnaires was processed with the use of SPSS® version 10.0 and presented with the use of tables. The qualitative data that were obtained from the interviews with senior government officers and archives personnel were content analyzed. Content analysis involved quantifying and tallying the presence of a concept. After identifying the categories, the data was coded. The coded data offered some evidence about dominant categories and trends. Some of the data was presented in narrative form or was integrated into the quantitative data by means of questionnaires.

5 Major findings and analysis of results

Starting with the nature of records management advice provided by KNADS, the following discussion presents major research findings and the analysis of the results.

5.1 Nature of record management advice provided by KNADS

The study findings indicated that all 157 (100%) registry personnel were aware of the existence of KNADS. When asked to indicate if they received professional records management advice from KNADS, all 157 (100%) indicated a "yes" response. Senior government officers were asked to state if they sought assistance and collaborated with

KNADS. All 10 (100%) indicated that they collaborated with, and sought assistance from, KNADS. Typical responses as to whether they sought assistance from KNADS were as follows:

> "...yes we do contact Kenya National Archives to assist in managing our records…"

> "...various government circulars from the Office of the President and Directorate of Personnel Management have been encouraging us to liaise with Kenya National Archives to improve recordkeeping in our ministries…"

Registry personnel were also asked to specify the nature of records management advice they received from KNADS and their responses varied, as indicated in Table 1.

TABLE 1: NATURE OF RECORDS MANAGEMENT ADVICE RECEIVED FROM KNADS ARCHIVES PERSONNEL (N=157)

Advice received	Frequency	Percentage
Registry management	144	91.7
Records preservation	127	80.9
Files management	113	72
Records security	106	67.5
Reports management	43	27.4
Forms management	43	27.4
Disaster management	37	23.6
Directives management	32	20.4

Table 1 indicates that the most cited records management advice received from KNADS was registry management with a score of 144 (91.7%) followed by records preservation with a score of 127 (80.9%). The least cited advice was on directives management and disaster management, accounting for 32 (20.4%) and 37 (23.6%) responses respectively. Furthermore, all the 157 (100%) respondents reported that they did not have a forms management programme. Various forms are used in government ministries (eg in recruitment, training needs assessment, claiming retirement benefits, accounting for use of resources and procurement of supplies and equipment). This meant that registries did not have a comprehensive records management programme. A complete records management programme includes a multitude of disciplines, including forms, reports and directives management (Penn, Pennix & Coulson 1994:5).

When asked about the means they used to contact KNADS, respondents gave different replies, as shown in Table 2.

TABLE 2: MEANS USED BY REGISTRY PERSONNEL TO CONTACT KNADS (N=157)

Means used	Frequency	Percentage
Correspondence	124	79
Personal visits	97	61.6
Telephone	71	45.2
E-mail	33	21
Fax	6	3.8

It is evident from Table 2 that the most cited means used to contact KNADS was corre-spondence, which accounted for 124 (79%) of the responses. The least cited means was the use of fax and e-mail, which accounted for six (3.8%) and 33 (21%) respectively.

Respondents were asked if they faced any problems when seeking professional advice from KNADS. One hundred and eight (68.8%) did not face any problems, while 49 (31.2%) indicated that they did face problems when seeking professional advice from KNADS. The 49 (31.2%) respondents who indicated that they had faced problems when seeking professional advice from KNADS were asked about the nature of these problems. Thirty-six (73.5%) cited delays in response, while 13 (26.5%) cited non-response from KNADS. All respondents pointed out that KNADS always avoided dealing with electronic records when asked for assistance by government departments. Registry personnel did not therefore see how they could influence the management of electronic records in their respective departments.

On their part, all six archive staff confirmed that they provided record management advice to government departments. They indicated that the areas of professional records management services that they provided to registry personnel in government depart-ments were: designing and implementing registry management policies and procedures, conducting records surveys and appraisals, preparing record retention schedules and records disposition and, finally, providing records management training to registry personnel and heads of departments.

Respondents were also asked to provide statistics indicating the nature of professional records management advice they had provided to government departments in the past five years. Three out of the six archives personnel provided the statistics indicated in Table 3. Although the statistics provided by the three archives personnel may not have reflected all the offices visited in the period concerned, they nevertheless provide an indication of some of the offices that were visited (and the advice that was given).

127

TABLE 3: RECORDS MANAGEMENT ADVICE PROVIDED BY KNADS TO DEPARTMENTS IN THE LAST FIVE YEARS

Advice	Ministry/Department	Year
Developing records management policy	Foreign Affairs	2004
	Finance (Kenya Revenue Authority)	
	Tourism and Wildlife	
	Energy (National Oil Corporation)	
	Trade (Kenya Bureau of Standards, Kenya National Trade Corporation)	2003
	Wildlife (Kenya Tourist Development Corporation)	
	Information and Communication (Kenya Institute of Mass Communication, Postal Corporation of Kenya)	
	Water and Irrigation (National Irrigation Board)	
	Agriculture (National Cereals and Produce Board, Kenya Tea Development Authority)	2002
	Labour and Housing (National Housing Corporation, Export Processing Zone, Investment Promotion Council)	
	Gender, Culture and Sports	
	Energy (Kenya Pipeline)	
Records classification	Justice and Constitutional Affairs (Governance and Ethics Dept)	2005
	Vice-President and Ministry of Home Affairs (Probation Dept)	
	Planning and National Development (HQ)	2004
	Office of the President (State House, Provincial Administration)	2003
	Finance (HQ)	2001
File indexes	Justice and Constitutional Affairs (Governance and Ethics Dept)	2005
	Vice-President and Ministry of Home Affairs (Probation Dept)	
Records survey and appraisal	Planning and National Development (HQ)	2005
	Planning and National Development (HQ)	2004
	Health (HQ)	
	Office of the President (State House, Provincial Administration)	2003
	Vice-President and Ministry of Home Affairs (Prisons Dept)	2002
	Finance (HQ)	2001

128

Records disposal	Labour (HQ)	2004
	Office of the President (Department of Defence)	
	Education, Science and Technology (Kenya Literature Bureau)	
	Energy (National Oil Corporation)	
	Trade (Kenya Bureau of Standards, Kenya National Trade Corporation)	2003
	Wildlife (Kenya Tourist Development Corporation)	
	Agriculture (National Cereals and Produce Board, Kenya Tea Development Authority)	2002
	Labour and Housing (National Housing Corporation, Export Processing Zone, Investment Promotion Council)	
	Gender, Culture and Sports	
	Energy (Kenya Pipeline)	
Training registry staff	Planning and National Development (HQ)	2004

Table 3 illustrates that archives personnel from Nairobi Records Centre provided records management advice in the last five years to various government departments and state corporations (parastatals) within various departments (eg the Kenya Tea Development Authority). Overall, the nature of advice provided included:
• records classification and index systems
• records survey and appraisal
• records disposition
• developing records management policy
• training registry personnel

Certain trends and patterns can be observed from Table 3 regarding the record management advice provided by KNADS to government departments in the last five years. For example: in the year 2001, few institutions were visited and the range of records management advice was limited to filing systems and records survey and appraisal. However, in the subsequent three years, that is, 2002 to 2004, the number of visits by KNADS archives personnel to government departments increased, as did the range of records management advice provided. For example, advice was provided in areas such as records disposal, developing records management policy, file classification and training of registry personnel.

The emphasis on records surveys and appraisal and records disposal by KNADS archives personnel in the years 2002, 2003, 2004 and 2005 may be attributed to the various records management circulars issued by the Office of the President and Directorate of Personnel Management. These circulars referred to the fact that many ministries had semi-current and non-current records mixed up with current records, thus making it impossible for registries to provide accurate records and the information needed for quick and timely decision-making.

129

The circulars advised government departments to seek the advice of KNADS in order to streamline their registry management practices as a way of improving public service delivery. For example, between 1999 and 2003, the government issued four record management circulars. The issuing of record management circulars may have led to government departments seeking more professional advice from KNADS on how to manage their records.

During the period 2001 to 2005, the Government of Kenya was carrying out public sector reform programmes and records management was identified as being one of the critical factors that would determine the success of the reform programmes. In this, government departments sought a great deal of assistance from KNADS. On the other hand, KNADS was keen to provide record management advice to government departments as a way of not only justifying its central role in the successful implementation of public sector reforms, but also as a way to market its services within the Kenyan public service.

All six archives personnel pointed out that they conducted records surveys. As far as the issues covered when conducting records surveys, are concerned, four out of six cited reviews of filing classification and indexing systems, types of records created/received, record storage conditions, developing record disposition schedules and reviewing record distribution systems. The remaining two respondents also cited training needs assessment in records management. When respondents were asked to indicate how frequently they conducted records surveys, their responses were as follows:
- " according to our work programme"
- "records surveys are not conducted regularly"
- "once every three to five years depending on the workload"
- "according to need in public offices vis-à-vis our resources"
- "as and when we complete one provenance (ministry/department)"

Records surveys are essential. They help national archives to manage records created by government throughout their life-cycle and to develop sustainable records management programmes. In other words, record surveys are the keys to establishing, maintaining and improving records management systems (Ngulube & Tafor 2006:62). It would be very difficult for KNADS to effectively monitor and improve record management practices in the public sector in the absence of a clearly defined records survey programme supported by adequate human resources. For instance, the size of the staff compliment influenced KNADS' capacity to deliver services such as records surveys and appraisal (ESARBICA 2003).

It was established that all six archives personnel conducted records appraisal and that they used the value and functional based appraisal approach. One respondent cited

"uniqueness of information contained in the records" as an additional appraisal criterion used by archive personnel. All six indicated that they provided assistance to registry staff on the development of retention schedules and disposal of redundant records.

It is evident from the foregoing that government departments had not been given advice on managing electronic records. Indeed, all six archive staff indicated that they had neither undertaken a survey to determine the amount of electronic records generated in the public service nor provided advice to registry staff on how to manage electronic records. They further said that they did not have an electronic records management programme for public sector records. All six archive staff provided multiple responses stating the challenges they faced in relation to the management of electronic records in government departments --- see Table 4.

TABLE 4: CHALLENGES FACED BY ARCHIVES PERSONNEL IN MANAGING ELECTRONIC RECORDS (N=6)

Challenge faced	Total
Lack of adequate staff	6
Lack of adequate information technology training	6
Lack of adequate financial resources	5
Not determining amount of electronic records created	5
Inadequacy of existing records and archives legislation	5
Defining the role of system administrators and managers	5

Table 4 shows that the most cited challenges were lack of adequate information tech-nology training and inadequate staff, with a score of six each. The next challenges mentioned were being unable to determine the amount of electronic records created, lack of adequate financial resources, inadequacy of existing records and archives leg-islation, and the undefined role of system administrators and managers (score of five). In short, one is obliged to conclude that KNADS is not playing any meaningful role in advising government departments on the management of electronic records at a time when government websites are proliferating, and government business increasingly being conducted electronically.

5.2 Implementation of KNADS' advice

Archives personnel were asked to indicate if registry staff implemented the advice that they received relating to the management of records. Four said they did not, while the remaining two recorded a "yes" response, but added that this varied from office to office. The four respondents who indicated that registry personnel did not implement the advice provided relating to the management of their records were asked to state reasons for non-compliance from a list of options provided. Their multiple responses are indicated in Table 5.

131

TABLE 5: REASONS FOR NON-COMPLIANCE WITH RECORDS MANAGEMENT ADVICE PROVIDED TO REGISTRY PERSONNEL (N=4)

Reasons for non-compliance	Frequency	Percentage
Lack of support from senior government officers	4	100
Lack of trained registry staff	3	75
Constant transfers of registry staff	3	75
Non-enforcement of the provisions of Cap 19	3	75
Public Archives Act Cap 19 not giving record creators more responsibility to manage their own records	3	75
Low priority accorded to records management in government departments	2	50
Lack of regular follow-ups by Nairobi Record Centre staff	2	50

Table 5 shows that the most cited reason was lack of support from senior government officers, while the least cited reasons were low priority accorded to records management by government departments and lack of regular follow-ups by Nairobi Record Centre staff. The registries staff agreed with archives personnel, and added that the other problems that led to non-compliance with KNADS' advice was constant transfer of staff and low morale of registry personnel. According to these respondents, morale in registries was low because, in their own words:

- "Registries are neglected units in the ministries."
- "Record management is a neglected area in the ministries."
- "Registry personnel are ignored in the ministries."
- "Recruitment and deployment of registry staff not based on experience and qualifications."
- "Registries are viewed as dumping grounds for problematic staff."
- "Few opportunities for training."
- "Poor scheme of service."
- "Health and safety of registry staff not taken into account."

Perhaps all this at least partly explains why the current state of record keeping in the Kenyan public sector is far from satisfactory, a state of affairs noted by Musembi (2004). He observed that the state of record management in many public service delivery departments was appalling because of irregular visits to departments by KNADS and the low morale of registry staff.

The problems faced by KNADS in managing government records were once common in Botswana. An investigation by Kenosi (1999:119-127) on record management practices in the Botswana public service found that the national archives faced problems in the

132

management of active, semi-active and non-active records in the public service. These included inadequate record centres, lack of blueprint for the management of electronic records and inadequate retention and disposal schedules.

Other problems that archival institutions face and which influences their programmes relate to insufficient funding caused by many government officials' failure to recognise records and archives as evidence. According to Millar (2004), governments in many parts of the world, and particularly in developing countries, perceive archives as relics of the past, that is, ephemeral materials kept for their historical value, but not as evidence of the rights and obligations of governments or citizens. As a result, there was often limited financial and organisational support for archives programmes, and national archives in many developing countries were marginalised, with insufficient resources, deficient physical facilities, and an inadequate infrastructure. Similarly, the lack of adequate resources has greatly undermined the work of KNADS.

The non-existence of concrete departmental policies regarding the manner in which registries are run, delegation of records management duties to junior officers who are not trained in records management, and a distinct lack of enthusiasm in record management by senior officers (whose portfolio includes the management of registries) was observed by Wamukoya (1988:7-8). The study finding which indicated that registries were not highly regarded in government departments also concurs with Millar's view (2004), who noted that there was a general lack of recognition of the importance of records as evidence. Millar also made the point that senior officials often tend to fail to recognise the need for, or value of, effective records programmes. The support of senior management of record management activities in Kenya may lead to increased funding and assist in the implementation of the record management advice provided by KNADS to government departments.

The need for senior management to support record management activities was emphasised by Makhura and Du Toit (2005:224). Discussing their experiences (based on a records management and information user behaviour study at South African National Parks), they observed that senior management needed to support the concept of a well-run record management programme and vigorously enforce a record management culture, thereby ensuring that all employees were familiar with agreed procedures for all types of records.

5.3   Challenges facing KNADS

Six senior government officers stated that they faced challenges in collaborating with KNADS, while the remaining four responded in the negative. The challenges they faced included inadequate advice on managing electronic records, inadequate retention and disposal schedules, slow response to requests for records surveys, delayed feedback

133

on records surveys and inconsistent record management advice. They attributed these problems to insufficient funding, lack of human resources and an inadequate legal framework. Archives staff concurred with senior government officers on the problems that affected their performance and added that the weakness of the Public Archives and Documentation Service Act Cap 19 also contributed to their woes. All these problems resulted in KNADS failing to adequately provide professional leadership and guidance on matters concerning the management of public sector records.

This partly explains why Mnjama (2003:91-101) pointed out that the factors that contributed to the poor state of record-keeping in Kenya not only related to problems within the public institutions, but also related to KNADS' failure as the main advisor to the government in the management of public records. However, any attempt to find sustainable solutions to the problems of managing public sector records in Kenya must involve KNADS (Wamukoya 1996).

In the absence of regular records management advice from KNADS, it would be difficult for public offices to do any of the following: create, classify and index, have in place efficient records access and use practices, ensure good records storage conditions, conduct environmental monitoring and control, devise disaster management and security plans, survey, appraise, and prepare retention schedules to guide records disposition (irrespective of format). It would also be difficult for archives personnel to make follow-up visits to government departments to monitor and ensure that the advice they provided to them was being implemented if human and financial resources are limited. Archives personnel would therefore not know the problems record creators faced in implementing the advice they provided and would therefore be unable to seek possible solutions to these problems.

Archives staff noted that the Public Archives and Documentation Service Act Cap 19 did not give record creators enough responsibility in managing their own records. Secondly, four out of six archives staff indicated that Cap 19 did not stress the continuity of records care during their entire life-cycle. Five out of all six archives personnel indicated that nor did the Act establish a partnership between record creating agencies and KNADS. The problem of record creators not implementing advice provided by national archives personnel is not unique to Kenya. In France, Barbat (1999) noted that establishing the responsibility of record creators was the most problematic area of record management. The improvements made in France since the end of the 1960s were due to the appointment of government ministry archive curators. Appointing officers in charge of record management by the national archives within departments was a worthwhile initiative, since it gave the archivists at national level someone to speak to in the government agencies (ie someone who knew about questions connected with records and who was responsible for the care of records).

134

In this regard, one way of addressing the problem of creators not implementing record management advice provided by archives personnel would be the recruitment and deployment of a record management cadre in government departments that was answerable to the National Archives. Other countries which have deployed record management officers in government ministries with specific record management responsibilities include India (The Public Records Act of India 1993) and Botswana (Chebani 2005; Mnjama 2005). For instance, Chebani (2005:139) indicated that Botswana National Archives and Records Service faced problems in managing records prior to the introduction of a records management cadre in the public service. Before this initiative, government records were managed by arbitrarily nominated administration staff who used inconsistent record management methods and practices.

Four out of two archives personnel indicated that the Public Archives and Documentation Service Act Cap 19 did not give the Director of KNADS authority to conduct research with a view to improving recordkeeping in the public service. They also said that the Act did not give the Director the mandate to manage electronic records or to collect and disseminate information on technological developments relating to record management in public agencies. An example of a technological development is the application of computers in the conduct of government business. On-going government e-governance initiatives in the public service would lead to increased use of computers in government departments. In other words, electronically generated records may be lost if registry staff are not kept informed of such developments.

6 Conclusions

Even though archives personnel provided record management advice to registry personnel in areas such file classification and indexing systems, records surveys and appraisal, records disposition, and developing a record management policy (see Table 3), the advice was not provided on a regular basis owing to the constraints they faced. One key finding of the study was that registry personnel indicated that they did not receive advice on reports and forms management 114 (72.6%), directives management 125 (79.6%) and disaster management 120 (76.4%). These findings show that KNADS need to diversify the scope of professional record management advice provided to registry staff to include these areas.

These findings suggest that, although KNADS played a key role in advising public agencies to manage their records effectively as mandated by the Public Archives and Documentation Services Act Cap 19, it had not sufficiently carried out this mandate. It can therefore be concluded that KNADS had not effectively helped registry staff to manage records, irrespective of their format and during their continuum, as a result

135

of resource constraints, and that this may have compromised public service delivery (since trustworthy decisions are based on well-managed information).

The advice provided by KNADS archives personnel did not cover other areas of record management such as the management of electronic records and disaster management. KNADS archives personnel possessed inadequate information technology skills. In fact, record management advice was fragmented and inconsistent. Registry staff did not implement advice that was provided relating to the management of records. Some of the reasons for non-compliance included constant transfers of registry staff, lack of trained registry staff, lack of support from senior government officers and public archive Act Cap 19, which did not give record creators enough responsibility in managing their own records. Furthermore, the Act did not provide the national archives with adequate powers for the management of electronic records.

However, all is not lost in Kenya in relation to the effective management of public re-cords. Mnjama (2005) noted that Kenya, along with Botswana, South Africa, Tanzania, Zanzibar and Zambia, was one of the archival institutions that had made considerable progress in managing non-current records. Kenya may build on the existing strengths in managing non-current records to effectively manage its current records.

7 Recommendations

Although it is commendable that KNADS continues to provide record management advice to record creating agencies to improve public service delivery, it is strongly recommended that this advice be provided on a regular and continuing basis. It is also recommended that the scope of records management advice provided to record creating agencies be broadened to include other components of a records management programme, namely: reports, forms and directives management, records preservation and security and disaster management.

Even though KNADS provided record management advice to record creating agencies such advice was not, in most cases, implemented, particularly as far as the disposal of valueless records was concerned. To ensure the implementation of record management advice provided to records creating agencies and enhance public service delivery, the researcher recommends that the KNADS lobby the government to appoint departmental record officers who will work closely with KNADS archives personnel, to improve recordkeeping policies and ensure implementation of the records management advice provided. In order to raise the profile of record management in the public service, staff in the registries should be recruited and deployed on the basis of their record management qualifications and experience. Furthermore, archives personnel should be

136

equipped with information technology skills to enable them provide advice to record creating agencies on how to manage electronic records.

The study recommends that KNADS review its current approach regarding its records management role in the Kenyan public sector. Given the KNADS limited financial and personnel resources, the researcher believes that there is a need to review existing archives legislation to provide record creating agencies with more responsibility in the management of their own records. This would improve the management of records and this, in turn, would lead to enhanced public service delivery. Giving records creating agencies more responsibility over their records also helps in the management of electronic records, since many national archival institutions do not have the infrastructure to manage electronic records.

The study also recommends that, in order to improve public service delivery and the management of records in departments, KNADS, in conjunction with the Office of the President and the Directorate of Personnel Management, needs to organise seminars and workshops to sensitise senior ministerial officers on the strategic importance of record management in public service delivery and good governance. Such a sensitisation programme should be aimed at encouraging senior officers to take a direct and personal interest in ensuring that the record management advice provided by archives personnel from KNADS is implemented. Thus the concern expressed by the Office of the President, Directorate of Personnel Management (2003b), that some top administrators in government departments did not understand or appreciate the vital role played by registries, and therefore did not give the registries the necessary support and guidance, would be at least be partially addressed.

The study also recommends that Cap 19 be reviewed, to give the Director of KNADS the power to ensure compliance with standards that apply to the management of public records. The Act should be reviewed to facilitate the sharing between government agencies and the National Archives of the responsibility for record-keeping, as is the case in the National Archives of Australia (2004). Archival legislation should also give KNADS a clear mandate to manage electronic records.

The study established that KNADS had not undertaken a survey to establish the growth of electronic records generated in the public service. It is recommended that research be undertaken to establish the amount of electronic records generated in the public service. Such a study would also need to establish the conditions under which electronic records are generated in the public service and investigate how the existing electronic record management situation in the public sector would impact on e-government initiatives. Additionally, similar research may also be conducted in other countries in Africa in order to establish the role that national archival institutions play in the management of records, and to obtain data for comparison purposes.

References

Barbat, P. 1999. Discovery of records management in France and its consequences. Paper presented at the 1999 Archives and Records Management Seminar in Melbourne, Australia. Available: http://www.caldeson.com/RIMOS/barbat.html (Accessed 13 May 2006).

Blake, R. 2005. National archives. Complying with the records management code: evaluation workbook and methodology. Consultation draft. Available: http://www.nationalarchives.gov.uk/news/pdf/record_management_compliance.pdf (Accessed 22 May 2005).

Chebani, B. 2005. Merits and challenges of the integrated records services in the public service – a case of Botswana. *ESARBICA Journal* 24:139-156.

ESARBICA**.** 2003. Minutes of the ESARBICA Executive Board Meeting held in Maputo, Mozambique, 22-23 July.

Government of Botswana Ministry of Youth, Sports and Culture. 2007. Botswana National Archives and Records Services. *Mission statement*. Available: http://www.gov.bw/index.php?option=com_content&task=view&id...> (Accessed 16 July 2007).

ICA. 2004. Welcome to ICA. Available: http://www.ica.org> (Accessed 7 September 2004).

Ilorah, R. 2004. NEPAD: the needs and obstacles. *African Development Review* 16 (2):223-251.

Kenosi, L. 1999. Records management in the public service and the Botswana's national archives and records service. *African Journal of Library, Archives and Information Science* 9(2):119-127.

Kenosi, L. 2005. Official secrecy and the management of security records in the global age of terrorism and information. *ESARBICA Journal* 24:45-51.

Kenya National Archives and Documentation Service. 2000. *Mission statement.* Available: http://www.kenyarchives.go.ke/mission.htm (Accessed12 December 2005).

Kenya National Archives and Documentation Service. 2000. A word from the director. Available: http://www.kenyarchives.go.ke/director-main.htm (Accessed12 December 2005).

Makhura, M & Du Toit, A**.** 2005. Records management and information user behaviour at SANParks: a case study. *Mousaion* 23(2):213-229.

Millar, L. 2004. *Authenticity of electronic records: a report prepared for UNESCO and the International Council on Archives*. Available: http://www.ica.org/biblio/study13_2Erev.pdf (Accessed23 May 2006.

Mnjama, N. 2003. Archives and records management in Kenya: problems and prospects. *Records Management Journal* 13(2):91-101.

Mnjama, N. 2005. Archival landscape in eastern and southern Africa. *Library Management* 26 (8/9):457-70.

Musembi, M. 2004. Introduction to records management. Paper read at the Directorate of Personnel Management Training Workshop for Registry Supervisors on Records Management, Machakos, 21-24 August.

National Archives of Australia. 2004. Commonwealth recordkeeping – an overview. Available: http://www.naa.gov.au/recordkeeping/overview/new-approach.html (Accessed 29 July 2004).

138

National Archives of Australia. 2004. Recordkeeping standards and advice. Available: http://www.naa.gov.au/about_us/about_us.html (Accessed 27 July 2004).

National Archives of India. 2005. Role in records management. Available: http://www.nationalarchives.nic.in/record_mgmt.html (Accessed 22 August 2005).

National Archives and Records Administration. 2004. What is a record? Available: http://www.archives.gov/records-management/ (Accessed 27 July 2004).

National Archives and Records Service of South Africa. 2004. About the National Archives and Records Service. Available: http://www.national.archives.gov.za.aboutnasa-content.html (Accessed 23 June2004).

National Archives UK. 2004. Records management. Available: http://www.nationalarchive.gov.uk/recordsmanageement/ (Accessed 27 July 2004).

Ngulube, P. 2005. Research procedures used by Master of Information Studies students at the University of Natal in the period 1982-2002 with special reference to their sampling techniques and survey response rates: a methodological discourse. *The International Information and Library Review* 37:127-143.

Ngulube, P & Tafor, V. 2006. An overview of the management of public records and archives in the member countries of the East and Southern Africa Regional Branch of the International Council on Archives (ESARBICA). *Journal of the Society of Archivists* 27(1):57-83.

Obudho, M. 1999. Speech by permanent secretary, Ministry of Home Affairs and National Heritage, during the official opening of the Kenya National Archives and Documentation Service Senior Staff Bi-annual Seminar. Nairobi, Kenya, 25-29 October.

Office of the President. 1999. *Permanent secretary to the cabinet and head of public service circular OP.39/2A on cases of missing and lost files and documents in the public service.* Nairobi: Office of the President.

Office of the President. Directorate of Personnel Management. 2003a. *Terms of reference for a task force to improve the performance of registries in government ministries/departments.* Nairobi: Office of the President.

Office of the President. Directorate of Personnel Management. 2003b. *Circular DPM.4/10A (9) on streamlining the performance of registries in government ministries, departments, local authorities, provinces and districts*. Nairobi: Office of the President.

Okumu, W. 2002. *The African renaissance: history, significance and strategy.* Trenton, NJ: Africa World Press.

Penn, I, Pennix, G & Coulson, J. 1994. *Records management handbook*. Aldershot: Gower.

The Public Archives and Documentation Service Act Cap 19 Laws of Kenya (Revised). 1991. Nairobi: Government Printer.

The Public Archives (Amendment Act), 2 of 1990. Nairobi: Government Printer.

The Public Records Act, 69 of India. 1993. Available: http://www.nationalarchives.nic.in/public-record93.html (Accessed 24 August 2005).

Shepherd, E & Yeo, G. 2003. *Managing records: a handbook of principles and practice*. London: Facet.

Terer, J. 2000. Cases of missing and lost files and documents in the public service. Press release. *East African Standard* 12 May: 1.

Thurston, A. 1996. Records management in Africa: old problems, dynamic new solutions. *Records Management Journal* 6(3):187-199.

United Nations. 2005. *The millennium development goals report*. Available: http://www.unstats.un.org/unsdd/mi/pdf/MDG%20Book.pdf (Accessed 21 May 2006).

Wamukoya, J. 1988. Significant steps towards the betterment of the service of the Kenya National Archives. *ESARBICA Journal* 10:6-12.

Wamukoya, J. 1996. Records management and administrative reform programmes in Kenya. PhD thesis, University College London.

140