South African Journal
of Information Management
[DoE accredited]

Peer Reviewed Article Vol.7(4) December 2005

e-Research support services: responding to a challenge facing the South African research and information communities

R.J. Page-Shipp
Pretoria, South Africa

M.M.P. Hammes
University of Pretoria
Pretoria, South Africa
monica.hammes@ais.up.ac.za

H. Pienaar
University of Pretoria
Pretoria, South Africa
heila.pienaar@ais.up.ac.za

F. Reagon
HSRC
Pretoria, South Africa
FReagon@hsrc.ac.za

G. Thomas
GAELIC
Johannesburg, South Africa
gwenda@fotim.ac.za

M.J. van Deventer
CSIR
Pretoria, South Africa
E-mail: mvandeve@csir.co.za

S. Veldsman
COSALC- SASLI
Pretoria, South Africa
sasli@cosalc.ac.za

Contents

Key words: e-Research, information services, national initiatives, national framework, South African Research Information Services, SARIS, open access, access to information

1 Introduction

The South African national research and development strategy was published in 2002. It invited all role players in the national innovation system to rethink their role and to find opportunities to face the challenge of increasing economic growth and improve the quality of life for all South Africans. It was clear that the strategy called for a renewal in the information services sector. It was anticipated that the strategy would require a level of information service support that was not available at any individual institution. One of the challenges, the declining affordability of access to the global research literature, was the initial spur for the establishment of the South African Research Information Services (SARIS) project team endeavour. A project team, rather than representative approach was taken to save cost and to ensure quick turnaround times.

The investigation soon revealed an emergence of a new paradigm, sometimes called e-Research, which presents a broader range of information service support challenges. The research goal was therefore adapted and the requirements for a national information service framework were investigated. e-Research, which is a family of activities, has emerged directly as a result of global connectivity, including the Internet, and of high performance computing. To participate effectively in and to advance e-Research will require that South Africa invests in technically complex and often costly infrastructural support services. As these are pre-competitive requirements and because there is great benefit to be obtained by sharing development and operating costs, a joint approach by the major players is desirable. Cooperation will result in:

- optimal utilization of scarce funding from the institutions involved;
- increased opportunities and capability to attract project funding from government and donors;
- improved coordination and use of scarce skills in the system; and
- reduced duplication of effort and costs.

Similarly, sharing the costs and infrastructure required to support information needs would result in huge benefits for all participants.

This article describes the outcome of qualitative research conducted by the SARIS project team. The research revealed that support and resources immediately available to South African researchers are not equitable because of the widely disparate resource base at the various institutions and in regions. The project team therefore decided to adopt the view of the researcher or research group rather than that of an institution. A variety of research methodologies were used. Published literature and information available via the Internet revealed that the international community has been grappling with the requirements for e-Research support for some time. A series of structured interviews, both nationally and internationally, followed. The research was concluded through a number of national workshops where the research findings were shared and further input was gathered.

The article commences with a breakdown of the components of e-Research and then describes a number of local and global trends that are promoting the development of e-Research. Finally it sets out a user-centric approach to the governance and management of an e-Research support service. The focus on the needs of the researcher has revealed a flexible generic solution that can be adapted to a wide range of situations.

top

2 e-Research perspective – providing the background for a coherent solution

The e-Research paradigm is a composite of three main trends:

1. The ability to transfer large volumes of data and to share computation capacity between remotely situated researchers is the basis of e-Science, *'faster, different, better research'* (T. Hey, Core Programme Director for UK eScience and Chair of the JISC committee, personal communication, 22 June 2004 – tony.hey@epsrc.ac.uk).
2. The need to make better use of expensively created databases by '*the active management and appraisal of data over the life cycle of scholarly and scientific interest'* is the basis of a new field of endeavour called digital curation (P. Burnhill, Acting Director, Digital Curation Centre, Edinburgh, UK, personal communication, 25 June 2004 – EDINA, University of Edinburgh, Hugh Robson Link Building, 15 George Square, Edinburgh, EH8 9XD).
3. Scholarly discourse now takes place on the dual playing field of commercial publication and open access.

An effective support service for these activities would need to embrace at least the following components:

- **Computation and data transmission infrastructure**. To support new or expanded scientific enterprises (in natural and social sciences as well as the humanities), the infrastructure should make it possible for researchers to transfer and otherwise share, between geographically distributed researchers or groups, large data-streams or datasets, including digital objects, and to share models and even computing capacity. The service would need to ensure ongoing operational efficiency and provide support to users without highly technical ICT skills.
- **Research tools and applications**. Many collaborative research activities require access to research tools and models that are held in geographically remote institutions. Finding the available tools and applications and making them visible and available to other researchers would require efficient support and the promotion of open source and open standards.
- **Willingness and ability to share primary e-Science data**. Making primary research data available to other researchers is part of the e-Science paradigm. One needs to store not only the database *per se*, but also sufficient metadata to enable a potential user to find relevant data and be satisfied with its value and provenance.
- **Scholarly discourse: commercial and open access to eContent**. The pursuit of high quality research requires ready access to the published work and data of other researchers, and the facility to publish one's own findings. Increasingly this takes place via the World-Wide Web. In the traditional commercial model, payment of a licence or other access fee is required, whereas an emerging model is of 'free' or 'open' access to online publications in formal journals or institutional or national repositories. The dynamics between these two systems are complex and it seems certain that they will co-exist for some time to come. Support services for commercially published sources are well established and relate to cost and access negotiations, licensing, copyright agreements and training. Open access is relatively new, and for it to achieve maturity, researchers will have to be both *willing* and *able* to use this mode. 'Willingness' involves issues such as reward and recognition, which should best be addressed outside the support structure. However, 'ability' relates to the provision of infrastructure that enables the potential of open access publication to be realized. Opportunities for support are obvious.
  Most academic recognition and reward systems are based on peer review of publications and citations, where commercial publications presently predominate. The adoption of open access does not imply the abandonment of the peer review principles, although practices may change somewhat. The open access process in itself provides further support opportunities.
  Emerging informal communication mechanisms (e.g. Web logging or blogs and RSS) are becoming widespread and also require support.
- **Digital curation and preservation**. This activity refers to the active management of datasets for their scientific and scholarly useful lifetimes, including the promotion of effective and widespread use. 'The past few decades in Science will come to be known as the Dark Ages because technology enabled the collection of so much data, but lack of support systems led to its subsequent loss' (Kuny 1997:3). Other ePrint archives, for example, theses and research reports, form part of this category. Support would require the setting of interoperability as well as content management standards. It would also require an ability to manage the expectations of a large number of stakeholders.
- **Innovation projects**. This is not a component of e-Research *per se*, but to ensure ongoing improvement in this rapidly developing scenario, provision must be made for innovation in technologies and applications to support the various processes. This would best be undertaken as funded pilot projects aimed at creating outputs that are useful to the e-Research community.

Chapters and whole books have been written on each of these topics. These relatively superficial descriptions are hopefully adequate for the purposes of this article.

From the researcher's point of view the data, information and tool components, listed above, have the following in common:

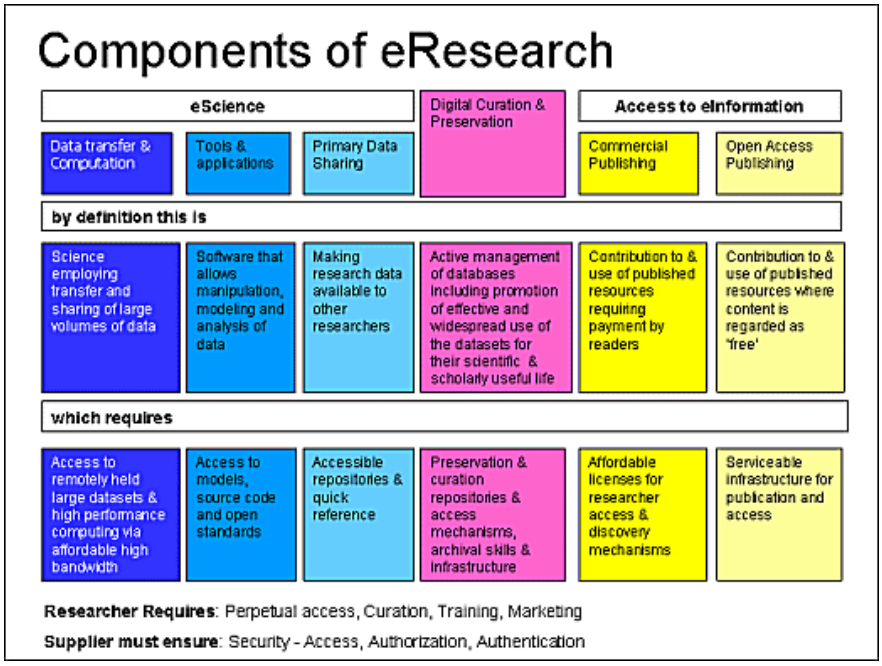
- Perpetual and pervasive access is required – items need to be available when and where they are needed.
- Long-term availability of output must be ensured. Most researchers view (with some justification based on recent experience) paper records as permanent and digital records as volatile. The support structures need to be trusted to maintain useful outputs after the immediate use on a project.
- Training has to be provided to optimize the usability of products and data sets.
- Marketing efforts are needed to ensure awareness of products and data sets.
- Security must be assured, so that only legitimate, authenticated users can have access, and then only for authorized usage.

The e-Research field is illustrated in Figure 1. The components set out above were grouped in two major subsets, e-Science and access to e-Information, with digital curation services fitting between them. Moving in from the left of the graphic one is concerned mainly with data, whereas textual information dominates from the right.

The establishment of an interface or service delivery mechanism, to gain easy access to the components, must of

course also be addressed. A Web services framework is the logical option, because it ensures that all researchers could gain access in a personalized manner. In well-resourced environments, the organization's portal could absorb the various components, which would be created in open standard format. Isolated researchers, or those in poorly resourced institutions, could gain access via a portal accessed from any Internet service point, be it a public library or an Internet café.

Figure 1 Components of e-Research



top

3 Alignment with current initiatives

Activities making up the family of e-Research are to be found in various stages of development in the research life of South Africa in 2005 but, typically, a 'Team South Africa' approach is not evident. One well-developed team initiative is the Coalition of South African Library Consortia's Site Licensing Initiative project (SASLI) that was created by the tertiary education institutions' library directors and to which the research councils have recently gained access. In a few short years, SASLI has made major inroads into reducing the costs of online access for researchers and has taken on several other relevant initiatives. The challenge relating to national site licensing is that a institution's ability to pay determines the sources to which the researcher has access. This policy often boils down to a feast or famine situation where researchers at well-resourced institutions have almost too much information and those in less well-resourced institutions have virtually none. A second challenge relates to sustainability. The SASLI activity, relative to what was later discovered internationally, is understaffed and has been dependent on donor funding. Other, sometimes distracting, initiatives have had to be launched to provide reciprocal services to donors or to create alternative funding streams.

Because of their novelty, e-Research activities have been characterized by embryonic, relatively small-scale initiatives which have served to build distributed expertise but also contributed to a dispersion of effort. Good use has often been made of overseas inputs but in many cases an initiative was a by-product of the efforts of a research team. Such initiatives make demands on research time and undermine the focus of the team who, unsurprisingly, have little concern for the cost-effective expansion of the services beyond their own needs, especially beyond termination of their research programme. Those groups that lack the expertise to create their own infrastructure are prevented from engaging in e-Research, possibly to the detriment of their visibility in the international research community. There is no coordinating body that is able to identify the entire set of ongoing activities. Asking for a 'coordinating body' should not be seen as a plea for centralization. A coherent, professionally managed, support system would ensure greater efficiency, with more rapid and cost-effective development of hardware and software solutions, while taking advantage of the distributed innovation capacity in the system.

Through the interaction with fellow researchers, the SARIS team members were able to identify significant national and global policy shifts and innovations that create an aligning framework for a coherent system. These include:

- the move of many commercially published scholarly journals and other text resources to digital format accessible via the Internet;
- the open access movement in scholarly discourse;
- various government initiatives aimed at ensuring that the results of publicly funded research are readily available to the public (House of Commons 2004; US House Appropriations Committee 2004);
- recent reports indicating new patterns in research and scientific communication and training, for example:
    - OCLC (2004) information format trends: content, not containers;
    - OCLC (2003) environmental scan: pattern recognition;
    - the Houghton (2003) report on changes in research patterns in Australia; and
    - a wide variety of items from the USA relating to the future of research and development in general and specifically to the research university (PCAST, The National Academies Project);
- widespread introduction of national and international research and education networks, providing low-cost, high-bandwidth communication between research and educational institutions [Internet 2 (USA), GéANT (Europe), AREN (Australia), Super Janet and the eScience Grid (UK) and, soon, South Africa's National Research and Education Network (SANReN)];
- digitization of important collections of objects and data, including documents;
- the emergence of digital curation as essential to ensuring that valuable data are preserved and made accessible for the duration of their scientific and scholarly useful life;
- the formation of international research networks such as the Human Genome Project and subsequent bio-informatics networks and the upcoming collaborations around the CERN Large Hadron Collider and the Square Kilometre Array Radio Telescope, all of which are critically dependent on rapid, cost-effective transfer of vast amounts of data; and
- the ICSU priority area study on scientific data and information (ICSU 2004)

Underlying all this is the, now conventional, use of the Internet as a means of near instant communication among researchers at a global level. Support services provide access to their services and products via conventional Web sites, portals or as entities within a Web services framework.

All of the above will impact on the way in which South Africa attempts to coordinate its e-Research activities.

A review of the above led to the formulation of a proposed framework for e-Research support services for South Africa, which was endorsed by the SARIS Advisory Group in December 2004 and has been expanded and redesigned somewhat in consultation with other stakeholders, in particular Dr Adi Paterson and his team at the South African Department of Science and Technology.

top

4 Framework for an e-Research services support system for South African researchers

If South Africa wants to sustain and expand its role as a player in the global research arena, we will have no choice but to coordinate our e-Research support systems to ensure cost-effectiveness and efficiency.

Key features would be:

- supporting world-class performance by South African researchers in the information society;
- as a pre-competitive activity, to be owned and driven by a passion for 'Team South Africa', with the top managers of the research delivery system, namely the tertiary education institutions and the science, engineering and technology institutions (SETIs) in the driving seat; and
- coordinating support activities to improve efficiency and not for the sake of consolidation.

Enlightened use of the 21st century potential for virtual enterprises should enable key activities to be housed in the most congenial place within the South African research system, rather than in a centralized agency. This would promote diversity and mitigate the evils of bureaucracy. In several countries where SARIS team members had made close acquaintance with fellow researchers (UK, Australia, Canada, Brazil and the USA), the creation of a professionally managed support service for e-Research activities has become a national or regional responsibility, but good performers throughout the system remain responsible for delivering the product or service that they are good at. The advantage is that the benefits from the enhancement of the total system reverberate to the whole community rather than to a single institution only.

Access to the support service should thus be Web-based and offer the researcher access to:

- a single, user-friendly, access point to a family of repositories for data, digital objects and publications that would act, *inter alia*, as a record of research outputs;
- easy communication with colleagues and other interested researchers;

- online content, including
    - a standard suite of online commercial information resources, paid for nationally;
    - other resources that are available without payment;
    - commercially published resources, in addition to the above, to which the researcher's institution subscribes; and
    - a pay-per-view facility for any other resources;
- assistance in finding, accessing and using databases that are needed for research;
- a global search engine that can search all those sources that are accessible to a particular researcher simultaneously;
- 'push' of references to published information – commercial and open access according to individual profiles;
- electronic communities of practice/curiosity;
- assistance in submitting large databases/streams to SANReN for transmission to colleagues and co-workers anywhere in the world, and in receiving such databases/streams; and
- online research support tools, for data and information/reference management.

The objectives of such a service would obviously need to be to:

- engage in the development of e-Research infrastructure that will establish standards, coordinate the innovative skills of the role-players in the system and distribute the benefits optimally to all participants; and
- manage investment in the development of the South African knowledge infrastructure for e-Research.

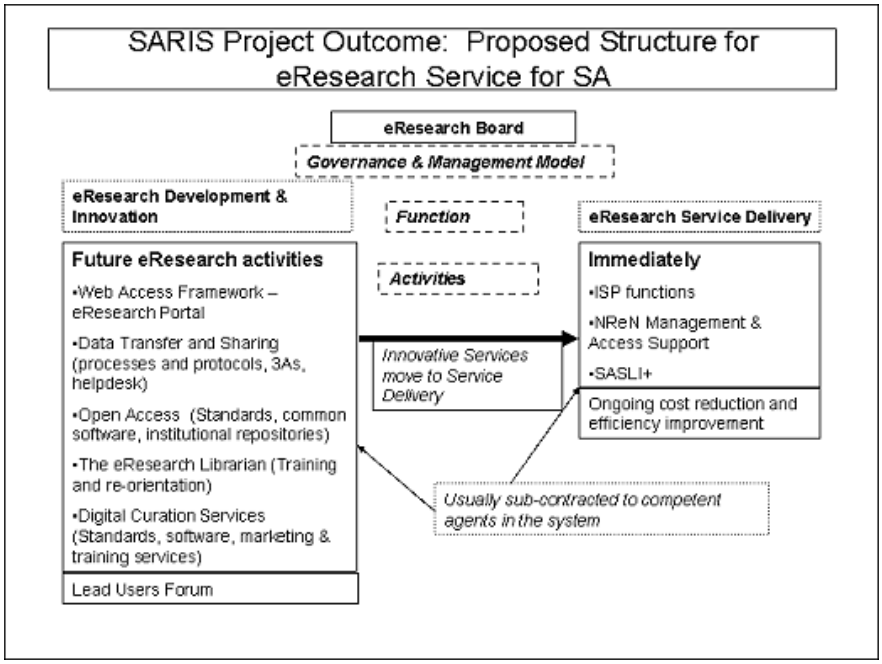
The efficient governance of a coherent system would be non-negotiable. It is essential that this should be a team effort at a high level in the research system. It is therefore proposed that an executive officer with a small core team, report to a representative management board. A governance structure is envisaged, involving the vice chancellors of tertiary education institutions and the chief executive officers of the SETIs (research councils), with active participation by the Department of Science and Technology (DST) and the Department of Education (DoE). A very similar structure already exists in the form of the tertiary and education network (TENET). This could mean that it may not be necessary to create a new entity because it would be more cost effective and in line with the coherent system approach to add this functionality to an existing entity rather than to propose a completely new one. It would, for example, be possible to strategically redefine the existing organization and to allow for a distributed network of content contributors.

Typical clients for a coherent system would be the government (especially the DOE and the DST, but also line departments with responsibility for research institutions), the top management of the tertiary education institutions and the research institutions, research deans and directors, directors of library and information services as well as information technology directors. The beneficiaries would be South African researchers, preferably across the entire system, including NGOs and possibly private sector institutions, although special financial arrangements may be necessary if the latter are to be included.

The proposal thus envisages a sophisticated and technologically advanced e-Research infrastructure that can manage, deliver, stimulate and reward. To develop and implement such a coherent infrastructure concurrently, starting from the current base, is unrealistic. It will be more practical to distinguish two groups of activity, namely e-Research innovation and e-Research service delivery. They would be linked in terms of accountability but distinct in roles and responsibilities. The framework shown in Figure 2 outlines such a managed e-Research infrastructure that meets these objectives and makes maximum use of existing entities, rather than creating new ones. It incorporates clear roles and a set of responsibilities that include:

- stimulating innovation and identifying appropriate innovation projects;
- obtaining project funding;
- feeding successfully completed projects into the service delivery component; and
- accountability to the South African research community

Figure 2 Proposed structure for e-Research support service for SA – a governance and management model



There is a grave danger that if the 'link' or interface between the two key components is left to function 'informally', the e-Research initiative could collapse. A governance structure must ensure that what is developed through innovation is delivered as services and support to the research communities where and when they are required. It is proposed that the e-Research board, consisting of specialist advisors, be constituted from among:

- DST, DoE
- TEI vice-chancellors
- SETI CEOs
- Other government departments
- SA Academy of Science
- Library directors
- IT directors
- Scientific, engineering and technological societies and allied professions group (SETAG).

There also needs to be an ongoing dialogue, and in some cases active collaboration, between the system developers and the service deliverers.

Responsibilities of the e-Research board would include the following:

- Devise policy direction for e-Research;
- drive innovation programmes for e-Research (strategic and business planning);
- identify projects, in consultation with relevant stakeholders, to advance innovation in e-Research;
- identify funding from appropriate sources for projects;
- promote e-Research at all appropriate levels in South Africa and internationally;
- migrate successful projects to the e-Research services delivery component and monitor their implementation; and
- ensure that the service delivery channels are maintained, upgraded, appropriately funded and deliver effectively.

An important feature of the framework is that neither the service delivery nor service innovation entities would necessarily undertake the development of identified innovations for the e-Research system, or their implementation. These activities should preferably be contracted out to competent players or groups of players within the research system who would create and maintain various service components, all of them accessible via the Web. It will also be possible to gain valuable support and shared learning via exchange agreements with the United Kingdom, Canada and Australian counterparts and other consortia networks (for example the USA's SOLINET) or national efforts (such as those at Brazil's IBICT).

top

5 Benefits

SA researchers and research groups, whether they function in and across well-resourced environments, are isolated (researchers in non-governmental organizations) or are participating in SADC/NEPAD-wide research groups, would benefit from having access to a coherent support service. The potential benefits are numerous but, most importantly, the service will enable and support those objectives of the national research and development strategy relating to an efficient science and technology system.

The availability of an effective e-Research infrastructure will not only support the country's top researchers but also assist researchers at all levels within the system to compete and maintain their profiles at a global level in their fields with minimal effort and time-wasting. In addition to the needs of full-time researchers, the systems should also support Government entities that can create value by the digital recording and analysis of operational data.

More specifically the system could benefit **researchers** by providing:

- improved, seamless and substantially equitable access to information for all South African researchers;
- the assurance that they could keep abreast of new products and content developments;
- training in the use of the products and services to:
    - ensure productive use;
    - considerably reduce the time spent searching for reliable content on the Internet; and
    - ensure that researchers have ongoing access to resources and are kept aware of their existence; and
- the infrastructure for more efficient knowledge generation.

Managers in the South African research system will benefit because:

- ownership of an essential research resource will be shared and the risk of losing the resource is therefore reduced significantly;
- the assurance that the potential afforded by the introduction of the SANReN can be exploited by the widest possible range of researchers, especially those whose conventional activities do not include the use of sophisticated information technology systems;
- increased focus on and visibility of the use of modelling, simulation and visualization in research will attract talented human capacity to South Africa;
- a mechanism for maintaining an holistic view of the South African e-Research activity will be maintained;
- reliable management information on costs and performance of systems providing access to scientific data and information will be available; and
- savings brought about by supporting more efficient operations could be utilized more effectively elsewhere in the research system.

Library directors and staff will also benefit.

- Better support systems will be available at a lower cost because the joint efforts of many role players will be pooled rather than duplicated.
- Exciting new challenges in the world of e-Research will surface previously ignored opportunities.

Most of all, 'Team South Africa' will benefit because there will be:

- heavyweight, national level representation during discussions with commercial providers of data and information, in some cases in an agency capacity; and
- sustainable and continued access to our own knowledge by providing standards for consistent curation and access processes effectively deploying and using curation services.

top

6 Conclusion

The research has revealed the potential for better positioning of the support needs of knowledge era researchers in the minds of top managers in the research system by drawing attention to the emerging phenomenon of e-Research and its particular support needs. The need for joint action by the participants if the e-Research challenges are to be met cost-effectively in South Africa is especially clear. It is proposed that a specialized agency should be created to provide the support that will fast track South Africa's contribution to global as well as locally relevant research efforts, and that the proposed governance model should work well for all the partners.

top

7 References

Houghton, J.W. 2003. *Changing research practices in the digital information and communication environment*. Department of Education, Science and Training, Australia.

House of Commons Science and Technology Committee report. 2004. *Scientific publications: free for all*. [Online]. Available WWW: http://www.biomedcentral.com/openaccess/inquiry/ (Accessed 17 March 2005).

International Council for Scientific Organizations (ICSU). 2004. *ICSU priority area assessment on scientific data and information*. [Online]. Available WWW: http://www.icsu.org/Gestion/img/ICSU_DOC_DOWNLOAD/551_DD_FILE_PAA_Data_and_Information.pdf (Accessed 17 March 2005).

Kuny, T. 1997. A digital dark ages? Challenges in the preservation of electronic information. *63RD IFLA Council and General Conference*. [Online]. Available WWW: http://www.ifla.org/IV/ifla63/63kuny1.pdf (Accessed 17 March 2005).

National Academies Project. *Impact of information technology on the research universities*. [Online]. Available WWW: http://www7.nationalacademies.org/itru/ITFRU_Workshop_Information.html (Accessed 17 March 2005).

OCLC. 2003. Environmental scan: pattern recognition. [Online] Available WWW: http://www.oclc.org/reports/2003escan.htm (Accessed 17 March 2005).

OCLC. 2004 Information format trends: content, not containers [Online] Available WWW: http://www.oclc.org/reports/2004format.htm (Accessed 17 March 2005).

PCAST. *President's Council of Advisorson Science and Technology. Reports and letters*. [Online]. Available WWW: http://www.ostp.gov/PCAST/pcast.html (Accessed 17 March 2005).

The Government of the Republic of South Africa. 2002. South Africa's national research and development strategy. [Online] Available WWW: http://www.dst.gov.za/legislation_policies/strategic_reps/sa_nat_rd_strat.htm (Accessed 17 March 2005).

US House Appropriations Committee. *National Institutes of Health report on open access*. 2004. NIH open-access plan frequently asked questions. [Online]. Available WWW: http://www.earlham.edu/~peters/fos/nihfaq.htm (Accessed 17 March 2005).

top

ISSN 1560-683X
Published by InterWord Communications for Department of Information and Knowledge Management, University of Johannesburg