

Towards Open Scholarship at the University of Pretoria

Theo Bothma^{*}, Heila Pienaar^{**} and Monica Hammes^{***}

^{}theo.bothma@up.ac.za*

University of Pretoria, Department of Information Science, Pretoria, South Africa,
<http://web.up.ac.za/default.asp?ipkCategoryID=3706>

*^{**}heila.pienaar@up.ac.za, ^{***}monica.hammes@up.ac.za*

University of Pretoria, Department of Library Services, Pretoria, South Africa

Abstract

World-wide the research paradigm is in the process of expanding into e-research and open scholarship. This implies new ways of collaboration, dissemination and reuse of research results, specifically via the Web. In the South African context the implication of e-research has been investigated at high level, and this resulted in the SARIS report (South African Research Information Services Report), according to which it is envisaged that South Africa, as a country, should position itself in the forefront of this new paradigm. This implies that individual research institutions should take the necessary steps to implement such strategies, collaborate amongst one another and lobby Government to support such initiatives.

At the University of Pretoria this new paradigm of e-research and open scholarship has led to the development of a high-level e-strategy for the University, supported by the Executive. This is embodied in different strategies, such an e-information strategy developed by the Department of Library Services (the Library, formerly known as the Academic Information Service), the Research Information System of the Research Office, and supported by a new IT strategy. This has resulted in a number of open scholarship projects, such as UPSpace (a digital research repository with a number of sub-projects, such as a repository of all UP research articles and Africana collections) and the UPeTD (a full-text database of theses and dissertations).

In this paper we give an overview of the rationale for the initiatives, briefly discuss the SARIS project and the e-strategies at the University of Pretoria and demonstrate the implementation of the various projects. We also discuss the technologies used, and reflect on the perceived advantages of the projects. The paper is illustrated by screen captures from different projects.

Introduction

Open scholarship, open-access scholarship and open-access scholarly publication all refer to approximately the same phenomenon, viz. the free and unrestricted access to scholarly publications via the Internet (see, e.g., Getz 2005a, 2005b, 2005c, Suber 2003, Bailey 2006). Based on the Budapest Open Access Initiative, the Bethesda Statement on Open Access Publishing and the Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities, Bailey (2006) lists the following characteristics of open access literature:

- It “is freely available”;
- It “is online”;

- It “is scholarly and royalty free”; and
- It “can be used with minimal restrictions”.

Bailey (2006) states that open access can be achieved through self-archiving and open access journals. Self-archiving typically occurs in open archives or open digital repositories, managed by research institutes or by universities. Clifford Lynch defines his view of an institutional repository as “...a university-based institutional repository is a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution” (Lynch, 2003).

This shift in scholarly publishing is accompanied by a transformation in research practice. “Research is becoming more multidisciplinary, more collaborative, and more global. [...] The term e-science has been used to describe large-scale, distributed, collaborative science enabled by the Internet and related technologies. E-research is a broader term that includes nonscientific research but that also refers to large-scale, distributed, national, or global collaboration in research. It typically entails harnessing the capacity of information and communication technology (ICT) systems.” (O’Brien, 2005).

South African Research Information Services (SARIS) project

A national research and development strategy for South Africa 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” (Page-Shipp et al., 2006). The SARIS project was started *inter alia* because of the extremely high costs to South African research institutes and university libraries to access the global research literature. From the research it was very soon clear that a new research paradigm, sometimes called e-research, was emerging and that this paradigm presented “a broader range of information service support challenges” (Page-Shipp et al., 2006). According to Page-Shipp et al. (2006) “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”.

The SARIS report (Page-Shipp et al., 2006) lists “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 [...];
- recent reports indicating new patterns in research and scientific communication and training, [...];
- widespread introduction of national and international research and education networks, providing low-cost, high-bandwidth communication between research and educational institutions [...];
- 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 [...]"

This led to a proposed structure for e-research support services for South Africa, as depicted in Figure 1.

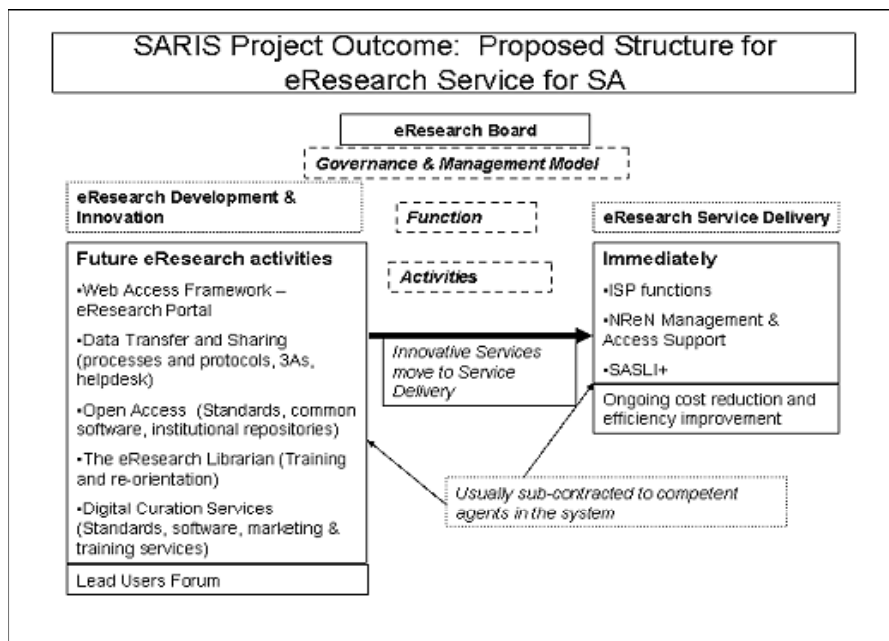


Figure 1: Proposed structure for e-Research support service for SA – a governance and management model (Page-Shipp et al. 2006).

The ideal was that the “future eResearch activities” of Figure 1 were to be coordinated at country level. However, it soon became evident that there would be no national coordination of these efforts in the near future, and that individual institutions should start their own initiatives.

E-strategies at the University of Pretoria (UP)

Since “becoming an internationally recognised South African teaching and research university is central to the University of Pretoria’s mission statement”¹ the University decided to align itself with the research activities of the SARIS project. This was captured in various strategic plans of the University and the Department of Library Services. The Strategic Plan of the University for 2007-2011, for example, states “[w]e believe it to be essential that the opportunities afforded us by these developments [= the development of new technologies]

¹ <http://web.up.ac.za/default.asp?ipkCategoryID=2995&subid=2995&ipklookid=2>.

should be fully exploited. We intend ensuring that this is the case” and “[t]he University of Pretoria will enhance the impact of its research and leverage the potential of its academic staff, students and networks of cooperation”.²

Based on this the Department of Library Services (also known as the Academic Information Service, AIS) developed their own e-strategy, which aims at the creation of an integrated seamless eService for the University of Pretoria. The objectives of the e-Strategy are (Pienaar 2006, Pienaar & Smith 2007a; see also Pienaar & Smith 2007b):

- To support education innovation and research excellence at UP;
- To deliver optimal e-information portal services (work flow) to our clients, and
- To take part in and make a contribution to international and national e-information phenomena, e.g. open access, digital preservation, e-Science, content management.

Key sub strategies were formulated in order to meet these objectives:

- The creation of an e-information environment for our clients;
- Development of an e-information plan as part of UP's e-strategy;
- Development of learning/e-learning & research/e-research support strategies;
- The adjustment of the AI's structure, business processes, skills and facilities to support the e-information strategy.

The e-information environment sub-strategy consists of the following projects: integrated systems, integrated interface, academic tools, global / federated searching, digital reference, ICT infrastructure, e-sources, e-dissertations, academic digital repositories, e-publication and digital preservation. The remainder of this paper focuses on two of these projects that make UP research visible on the web, viz. the UP digital repository, with a number of collections, such as a repository of all UP research articles and special collections and the UP electronic theses and dissertations database.

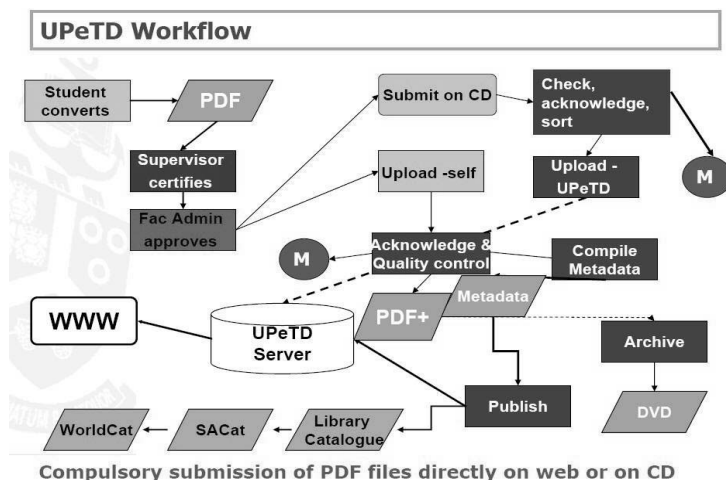


Figure 2: UPeTD workflow, from Hammes 2005.

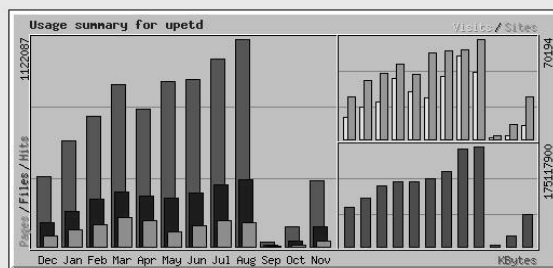
² <http://web.up.ac.za/default.asp?ipkCategoryID=2995&subid=2995&ipklookid=2>.

UPeTD – the UP electronic theses and dissertations database

The University decided in 2000 to start with a pilot project to make theses and dissertations available online. “Our theses and dissertations are proof of excellent research work done at high cost for the university as well as for the country: they deserve to be read widely” (Hammes 2005). Within the IT framework of the University of Pretoria it was decided to use open source software and open standards. At the time the ETD-db software developed by researchers at Virginia Tech was the only real solution for theses and dissertations and was chosen as platform. It is freely available through the NDLTD, the Networked Digital Library of Theses and Dissertations, an “international organization dedicated to promoting the adoption, creation, use, dissemination and preservation of electronic analogues to the traditional paper-based theses and dissertations”³. In the near future the database will migrate to one of the newer platforms.

Usage Statistics for upetd

Summary Period: Last 12 Months
Generated 11-Nov-2007 06:47 SAST



Summary by Month										
Month	Daily Avg				Monthly Totals					
	Hits	Files	Pages	Visits	Sites	KBytes	Visits	Pages	Files	Hits
Nov 2007	32377	9874	2941	872	29720	55579094	9602	32357	108624	356150
Oct 2007	35457	10169	2720	753	10122	18249577	2260	8161	30508	106373
Sep 2007	12346	3740	2028	484	2535	3309563	969	4057	7480	24693
Aug 2007	36196	11690	4151	1515	70194	175117900	46982	128692	362416	1122087
Jul 2007	32693	10765	4537	1892	62947	170676073	58664	140651	333739	1013491
Jun 2007	30121	9703	3825	1479	61613	131289916	44383	114754	291104	903643
May 2007	37310	11013	3337	1211	60446	118810750	29079	80109	264324	895449
Apr 2006	24784	9163	4612	1113	45813	113275718	33407	138389	274912	743548
Mar 2006	28224	9495	5024	1383	52532	112958933	42885	155763	294373	874952
Feb 2006	25254	9162	4317	938	45988	106226896	26274	120900	256558	707121
Jan 2006	18465	6195	2872	722	40863	85521449	22406	89058	192065	572437
Dec 2005	14540	4967	2141	589	29772	69309013	15329	55691	129154	378030
Totals						1160524882	332240	1068582	2545257	7697994

Figure 3: Usage statistics for UPeTD for the past twelve months, from <http://upetd.up.ac.za/webalizer/> (low figures for September and October are because of system problems).

³ <http://www.ndltd.org/>; details of the initiative can be found at <http://www.ndltd.org/info/index.en.html>.

The required policies and procedures were developed, and detailed workflow was documented, as indicated in Figure 2. In 2003 the first manuscripts were uploaded. Initially submission was voluntary, but a decision was taken by Senate that all students registered for a Masters or a doctorate from January 2004 had to submit their research in electronic format, either directly to the repository or on CD to Faculty Administration, before graduation. Minimum specifications for ETDs were set⁴ and detailed guidelines and tutorials were made available to students to simplify the process⁵.

Titles and abstracts of all theses and dissertations at the University of Pretoria have retrospectively been added to the database, even though not all theses and dissertations are available in full-text. Currently there are 3 226 titles in the database, of which 3 108 are available in full-text. Once the data has been submitted to the database it is automatically harvested by various harvesters, inter alia Google Scholar. It is evident that the UPeTD has made theses and dissertations at the University of Pretoria much more accessible and widely read, as is evidenced by the statistics in Figure 3.

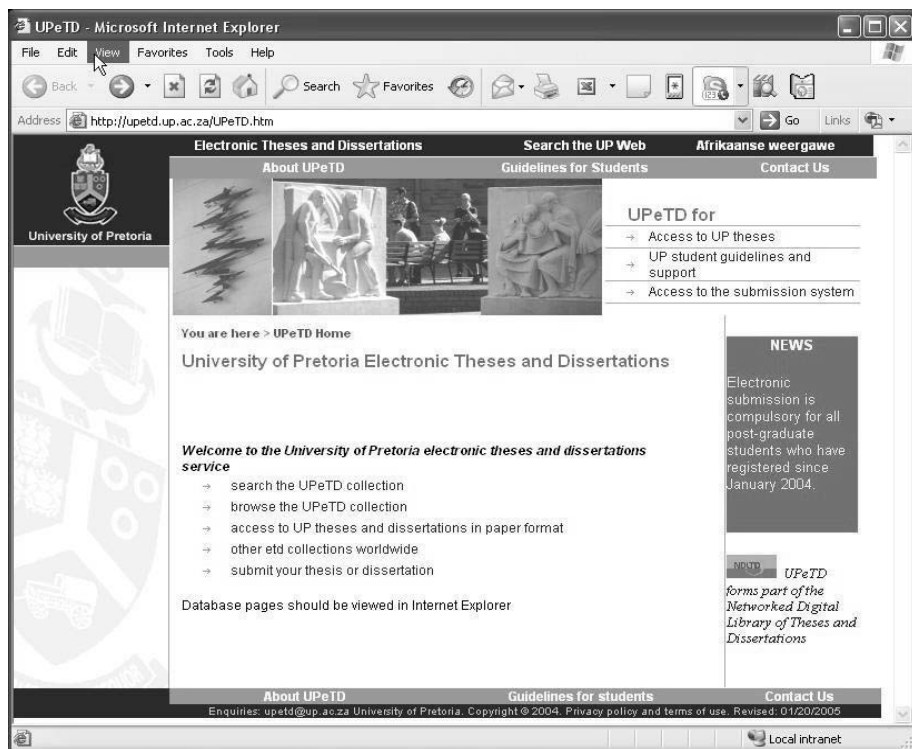


Figure 4: The home page of UPeTD at <http://upetd.up.ac.za/UPeTD.htm>.

⁴ <http://upetd.up.ac.za/authors/publish/standards.htm>.

⁵ <http://upetd.up.ac.za/authors/publish/index.htm>.

Figure 4 provides a screen capture of the home page of UpeTD⁶ and Figure 5 is a typical example of the title page for an eTD.

A recent survey of doctorates whose theses are available in the database indicated a number of very positive outcomes for their careers and scholarly collaboration (Hammes and Mahlangu 2007).

UPSpace – the UP digital research repository

The UP digital repository project team evaluated several open source and commercial software platforms, e.g. Greenstone, Innovative, Fedora, E-prints, DSpace, I-Tor over a six month period in 2004. DSpace was chosen because it fits the UP IT architecture and supports a distributed approach to an institutional digital repository; a further key issue is that it is open source. DSpace was developed by MIT and Hewlett Packard to address the preservation and dissemination needs of MIT. DSpace has at its core a model which attempts to address the real needs of a flexible institutional repository. This “Communities and Collections” structure aims to allow devolved control and administration from a central service to individual departments. DSpace is written in Java and utilises a PostgreSQL database layer. Long term digital preservation is a key aim of the DSpace system and MIT (and others) are investing considerable research in enhancing this aspect of the repository software. For details on DSpace see <http://www.dspace.org/>.

Document Type	Doctoral Thesis						
Author	Korsten, Hilda						
URN	etd-10252006-162402						
Document Title	An evaluation of and a model for South African government websites						
Degree	DPhil (Information Science)						
Department	Information Science						
Supervisor	Advisor Name	Title					
	Prof T J D Bothma						
Keywords	<ul style="list-style-type: none">• Web sites evaluation South Africa• Internet in public administration South Africa• Government communication systems South Africa• Internet in public relations South Africa						
Date	2002-12-01						
Availability	unrestricted						
Abstract							
Please read the abstract in the 00front part of this document							
Files							
Filename		Size	Approximate Download Time (Hours:Minutes:Seconds)				
			28.8 Modem	56K Modem	ISDN (64 Kb)	ISDN (128 Kb)	Higher-speed Access
00front.pdf		3.17 Mb	00:14:40	00:07:32	00:06:36	00:03:18	00:00:16
01chapter1.pdf		2.91 Mb	00:13:27	00:06:55	00:06:03	00:03:01	00:00:15
02chapter2.pdf		3.38 Mb	00:15:38	00:08:02	00:07:02	00:03:31	00:00:18
03chapter3.pdf		4.23 Mb	00:19:34	00:10:04	00:08:48	00:04:24	00:00:22

Figure 5: A typical example of the title page for an eTD.

⁶ <http://upetd.up.ac.za/UPeTD.htm>.

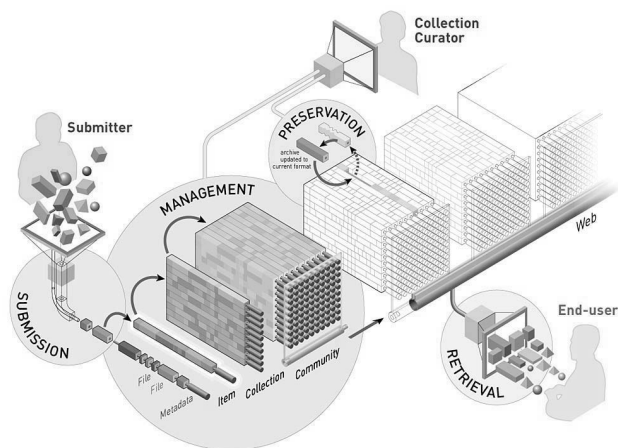


Figure 6: The DSpace digital repository model⁷.

The DSpace digital repository model is illustrated in Figure 6. DSpace is modular and therefore supports the creation of different repositories, even across institutional boundaries.⁸

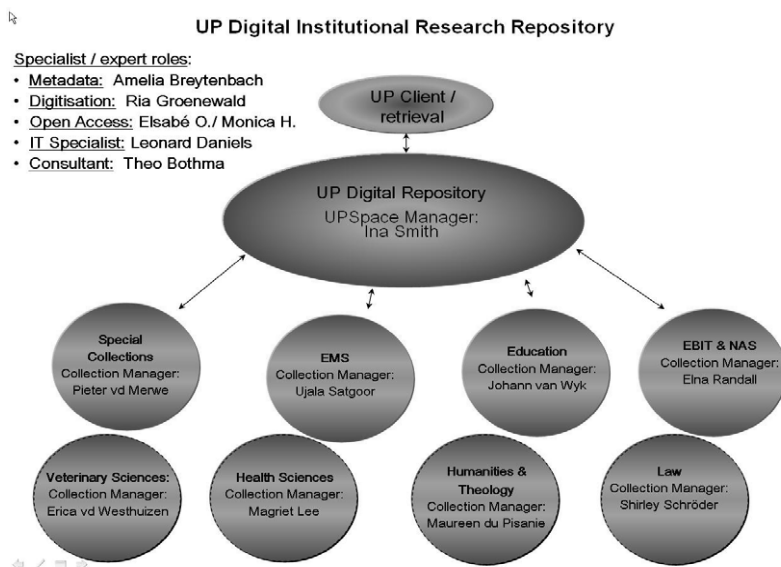


Figure 7: UP digital research repository management team.

⁷ <http://www.dspace.org/introduction/dspace-diagram.pdf>.

⁸ <http://www.dspace.org/introduction/dspace-diagram.pdf>.

DSpace was adapted in 2005 to support UP requirements, e.g. authentication via the UP portal and extensions to the metadata. A UP digital repository management team was created with specific areas of responsibility; see Figure 7. Specialist roles were identified for metadata, digitisation, platform manager, IT and an external consultant. Collection managers were appointed for different focus areas, e.g. scholarly communication (including e-prints), special and Africana collections, Veterinary Science Faculty, Department of Architecture, Education Faculty and the Mapungubwe museum. The collection of scholarly articles is known as openUP, and is an overlay collection of the larger UPSpace collection.

At the time of writing 2,845 full-text items have been uploaded to UPSpace.

Figures 8-10 provide screen captures of examples from the different collections, viz. the general homepage of UPSpace (Figure 8), a research article in openUP (Figure 9), and a compilation of images from various cultural collections in UPSpace (Figure 10).

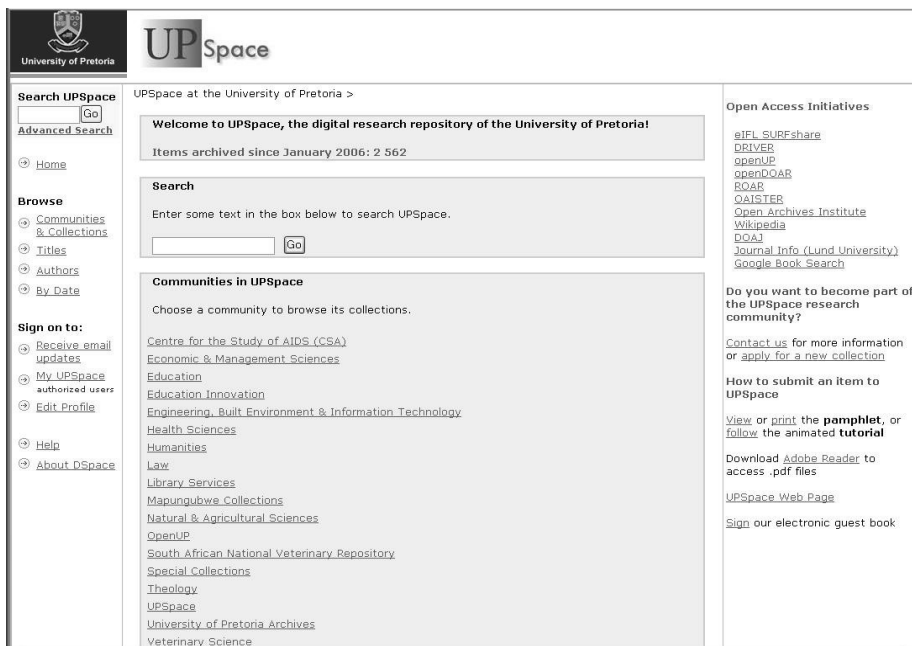


Figure 8: The homepage of UPSpace (<https://www.up.ac.za/dspace/>).

The ideal is to create an integrated view of UP research and researchers, by linking researchers' CVs, the University's annual research report and the full-text articles in openUP:

- Researchers can, at present, provide links from their CVs on their personal and/or departmental web pages to their full-text articles in openUP.
- The Research Office of the University is tasked to compile a full bibliographic list of all research at the University, and this is published in the Research Report; these

annual reports are available online⁹. The Research Office currently provides a link to all the openUP research article collections¹⁰. At present there are no links between the bibliographic listings in the annual research reports, but it is envisaged that this will be in place for the 2008 report.

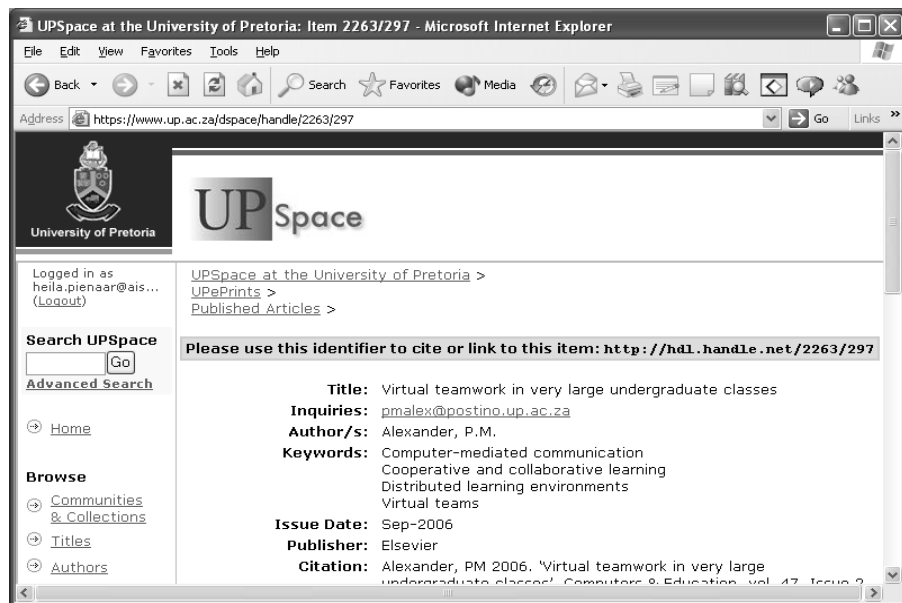


Figure 9: openUP – example of a published article.

Initially researchers were very hesitant to have their research archived in the UP repository. Many reasons were offered for this, which include issues about copyright and the extra work it entails. However, more and more researchers are being convinced that the effort is worthwhile because their research is becoming much more visible to the international scholarly community. Top management at the University has also been convinced of the value of the digital repositories, and a strategy is being developed to make submission of full-text articles in the repository compulsory, as in the case of theses and dissertations.

The creation of the various digital repositories has had a huge influence on the working lives of many librarians / information specialists at the University of Pretoria. Many had to be retrained (or had to restrain themselves) to adapt to the new environment, to become collection managers, digitization specialists, metadata specialists, open access managers, etc. There are many new challenges for information specialists as well, which include the following (Smith and Pienaar 2007); they have to:

- support the strategies and objectives of the institutional repository projects;
- influence the mindsets of colleagues in the library and also of researchers that may be hesitant to accept these new ideas;

⁹ <http://web.up.ac.za/default.asp?ipkCategoryID=1630>.

¹⁰ <http://web.up.ac.za/default.asp?ipkCategoryID=5280&subid=5280&ipklookid=14->.

- encourage knowledge transfer;
- motivate others to share, learn, apply their new knowledge;
- communicate new ideas and new ways of working effectively to their colleagues;
- identify new opportunities; and
- mentor others that may be struggling in the new work environment.

Examples UPSpace Content

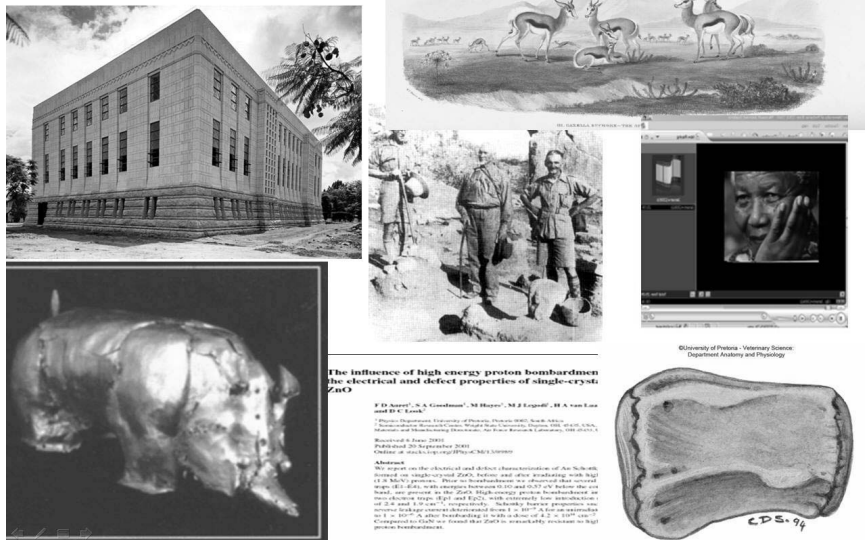


Figure 10: A compilation of images from various cultural collections in UPSpace.

A number of communities of practice emerged spontaneously to share the required knowledge to work effectively in the new environment. This sharing of knowledge is happening across institutional boundaries as well: members of the UPSpace team worked closely with information specialists from the CSIR (Council for Scientific and Industrial Research in Pretoria) to help them develop a digital repository as well; this is documented in a presentation at the IFLA Knowledge Management Workshop in Durban in 2007 (Pienaar and Van Deventer, 2007). The research team at the CSIR has, however, not only learned from the UP team – they have been able to provide a number of tips to the UP team to enhance their effectiveness.

The expertise that the UP team has obtained from their work during the past few years is starting to have a wider influence as well. Representatives from the Southern African Regional Universities Association (SARUA, an umbrella organization for universities in the southern part of Africa) visited UP in October 2007; they were most impressed with what the team at UP has accomplished and they are looking forward to working with UP on the

development of institutional repositories at all universities in the region. The presentation that they attended was delivered by Ina Smith and is available online¹¹ (Smith 2007).

A mailing list was also created on which many members from African countries are already registered. The e-mail address for this mailing list is irspace@kendy.up.ac.za. This is an informal mailing list for African and South African institutions with a common interest in institutional repositories, hosted by the University of Pretoria. Through this mailing list individuals are invited to collaborate, to share ideas, find solutions, and build innovations on institutional repositories.

Many additional initiatives for sharing knowledge about institutional repositories are envisaged. This includes the creation of a web page called “The Institutional Repository Toolbox” which will guide information specialists to set up institutional repositories in their own organizations. A wiki with references and links to articles on institutional repositories is also envisaged. Although there are not many institutional repositories yet in South Africa, there is already an entry on Wikipedia for South Africa’s institutional repositories at http://en.wikipedia.org/wiki/South_African_Institutional_Repositories.

Conclusion

Initially it took a huge amount of time and effort to start the various institutional repositories at the University of Pretoria. Many people had to be convinced that these are worthwhile efforts, inter alia top management. Hardware (scanners, servers), software had to be evaluated and bought / installed, policies and procedures had to be put in place; information professionals had to be retrained; the mindsets of researchers (and a number of information professionals!) had to be changed. And then the real work of populating the repositories started. As the collections grew the enthusiasm for the digital repositories also grew, and top management realized that their support for the endeavours was justified. More and more researchers are becoming convinced that open access and open scholarship is to their advantage. The skills that the information professionals at the University of Pretoria have acquired have empowered them to work effectively and efficiently in this new digital environment. They have been able to attain a very high level of specialization, to such an extent that their advice is being sought by other institutions. Through communities of practice, listservers, presentations, personal interaction and consultation they have been able to share their knowledge and expertise with information specialists from many institutions in South Africa and further afield.

Through our involvement in these projects for the past number of years we are convinced that the hard work and efforts that setting up the repositories required was worthwhile, and that:

- support from top management is essential to make a success of designing and implementing such a huge project;
- UP research has become much more visible to the rest of the world. Dissertations and research articles are regularly downloaded;
- researchers were originally rather reluctant to participate in the different projects, but that they have become much more enthusiastic once they realized that their research is being accessed by a much wider international audience;
- information professionals feel empowered through the process of re-training. Various communities of practice have been established to share knowledge; and that

¹¹ <https://www.up.ac.za/dspace/handle/2263/3810>.

- the skills and expertise of the information professionals at UP are at such a level of professionalism that they are being consulted by other institutions to help them set up their own repositories.

References

- Bailey, C. W. (2006). *What is open access?* Preprint: 2/7/06. Retrieved October 2007 from <http://www.digital-scholarship.org/cwb/WhatIsOA.pdf>.
- Getz, M. (2005a). *Three frontiers in open-access scholarship*. January 13, 2005. Retrieved October 2007 from <http://ecommons.library.cornell.edu/bitstream/1813/307/1/MGetz13Jan05Cornell.pdf>.
- Getz, M. (2005b). Open-access scholarly publishing in economic perspective. *Journal of Library Administration* 42/1, 1-39. Retrieved October 2007 from <http://www.haworthpress.com/web/JLA>.
- Getz, M. (2005c). *Open Scholarship and research universities*. Working paper No, 05-W17. Department of Economics, Vanderbilt University. Retrieved October 2007 from <http://www.vanderbilt.edu/Econ/wparchive/workpaper/vu05-w17.pdf>.
- Hammes, M. M. P. (2005). *Policies, standards & procedures for an ETD programme*. UWC ETD workshop, 2 June 2005. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/1441>.
- Hammes, M. M. P & Mahlangu, A. (2007). E-quality for equality. Paper presented at the *7th Northumbria Conference*, August 2007, Stellenbosch, South Africa. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/3405>.
- Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *ARL: A bimonthly report*, no. 226. Retrieved October 2007 from <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>.
- O'Brien, L. (2005). E-research: An imperative for strengthening institutional partnerships. *EDUCAUSE Review*, 40/6, 64–77. Retrieved October 2007 from <http://www.educause.edu/apps/er/erm05/erm0563.asp>.
- Page-Shipp, R. J., Hammes, M.M.P, Pienaar, H, Reagon, F, Thomas, G, Van Deventer, M.J. & Veldsman, S. (2005). e-Research support services: responding to a challenge facing the South African research and information communities, *South African Journal of Information Management*, 7/4. Retrieved October 2007 from <http://www.sajim.co.za/>.
- Pienaar, H. (2006). e-Information strategy of the Academic Information Service. Paper presented at the *DITCHE IT Conference & Workshop*, Port Elizabeth, 16-18 May 2006. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/1346>.
- Pienaar, H. & Smith, I. (2007a). Development and implementation of an e-information strategy for a university library. Paper presented at the *2nd Africa Libraries & Information Management Systems Conference 2007*, The Grace Hotel, Johannesburg (7 February 2007). Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/1681>.
- Pienaar, H. & Smith, I. (2007b). Development of a Library 2.0 service model for an African library. Presented during the *Third LIASA Gauteng North Branch Meeting*, 23 October 2007. IFLA feedback session. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/3779>.
- Pienaar, H. & Van Deventer, M. (2007). Capturing knowledge in institutional repositories ... playing leapfrog with giraffes. Paper presented at the *Knowledge Management pre-conference workshop* at the World Library and Information Congress (WLIC): 73rd IFLA General Conference and Council, 17 August 2007. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/3441>.

- Smith, I. (2007). Research @ universities: cited more, safe forever. Paper presented to the *Southern African Regional Universities Association*, 26 October 2007. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/3810>.
- Smith, I. & Pienaar, H. (2007). "Adapt or die"; Energizing library staff and academics through the development of digital repositories. Paper presented at the *World Library and Information Congress* (WLIC): 73rd IFLA General Conference and Council, 21 August 2007. Knowledge Management Section. Retrieved November 2007 from <https://www.up.ac.za/dspace/handle/2263/3325>.
- Suber, P. (2003). Removing barriers to research: An introduction to open access for librarians. *C&RL News*, 64/2. Retrieved October 2007 from <http://news.ala.org/ala/acrl/acrlpubs/crlnews/backissues/2003/february1/removingbarriers.cfm>.