Towards a digital preservation policy

Presented by Ria Groenewald & Ina Smith National Library IT Event

20 May 2008



Agenda



Part I	Digitization & Preservation	Ria Groenewald ria.groenewald@up.ac.za
Part II	Preservation & Trusted Digital Repositories	Ina Smith ina.smith@up.ac.za

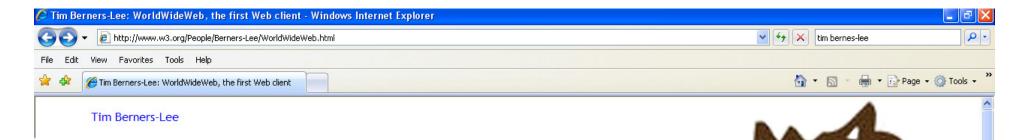




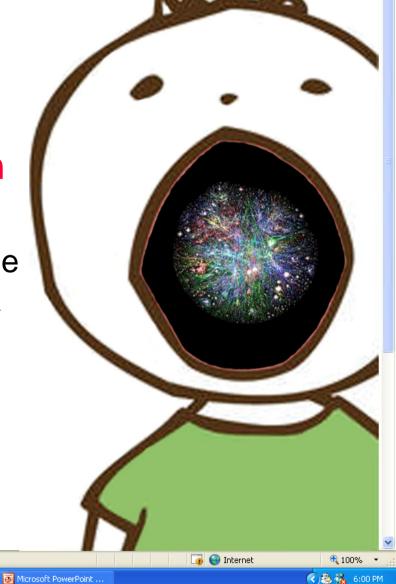
Part I: Digitization & Preservation

Ria Groenewald





- The birth of the web site 6 August 1991 - 2:56:20 pm
- The world's first website was made available on the public internet - a creation of Tim Bernes-Lee at **CERN**



8

DITCHE

Tim Berners-Lee: WorldWideWeb, the first Web client - Windows Internet Explorer	
The second se	V 🐓 🗙 tim bernes-lee
e Edit View Favorites Tools Help	
🔗 🌮 Tim Berners-Lee: WorldWideWeb, the first Web client	🟠 🔻 🔝 🝸 🖶 Page 👻 🎯 Tools 🔻
Tim Berners-Lee The WorldWide	Web browser
	s, after all, when it was written in 1990 it was the only way to see the web. Much ne abstract information space (which is now spelled <i>World Wide Web</i> with spaces).
	were some great tools available -it was a great computing environment in general.

In fact, I could do in a couple of months what would take more like a year on other platforms, because on the NeXT, a lot of it was done for me already. There was an application builder to make all the menus as quickly as you could dream them up. there were all the software parts to make a wysiwyg (what you see is what you get - in other words direct manipulation of text on screen as on the printed - or browsed page) word processor. I just had to add hypertext, (by subclassing the Text object)

This is a (242kB) screen shot of the browser, taken when things had got to the point that *Communications of the ACM* was interested in an article, in 1993. The differences between this and the first edition (Christmas 1990) were:

- The whole thing would have been grey scale as NeXTs were at the time just grey scale;
- The inline images such as the world/book icon and the CERN icon, would have been displayed in separate windows, as it didn't at first do inline images.

See also:

• another screen shot, this one by JFG, later but grey scale

A quick tour of this screen to answer the FAQs:

In this shot I am making a link from the word "ATLAS" in the list of experiments to some web page.

The NeXTStep operating system put the menu for each application in the top left of the screen. The application is called WorldWideWeb. because the menus are in this block they windows are very unencumbered. A little like like the windows "start" menu later.

The Navigate menu had things like "back" and "next" and "previous". these last two were useful when you follows a link from a list of links- they meant "go back a step and then take the next link from the same page instead".

The document menu was like the "file" menu for windows I suppose. The "find" menu is fairly self-explanatory, as is "edit".

関 Adobe Photoshop Alb...

The "Link" menu you can see. "Mark all" would remember the URI of where you were. "MArk selection" would make an anchor (link target) for the selected text,

🛃 start	💋 5 Internet Explorer	- 6.	
---------	-----------------------	------	--

DI:\DITCHE

🤇 The World Wide Web project - Windows Internet Explorer		
S S + E http://www.w3.org/History/19921103-hypertext/hypertext/WWW/TheProject.html	💌 🗲 🔀 world first web page	P-
🔁 -		
Search the Internet - type your phrase here 🔽 🏀 Enter 🔹 🛱 Block PopUp 🛛 🥒 🦻 🖏 🖓 Shop Ebay 🐗 Learn How to Play Poker 🔹 🚼 T	n Neue Haadiinaa 🦀 Calabiika Caasiin	
Search the Internet - type your phrase here 🔽 🏀 Enter 🔹 🗗 Block PopUp 🖉 🤌 獤 🖞 Shop Ebay 🗿 Learn How to Play Poker + 🚡 T Coogle 💽 world first web page 🔹 Go 🍬 🌍 🧭 🖕 🙀 Bookmarks + 🔊 2 blocked 🖓 Check + 🔦 AutoLink + 📔 AutoFill 🔒 Send to +	· · · · · · · · · · · · · · · · · · ·	🔘 Settings v
A A C The World Wide Web project	🙆 • 🗟 · 🖶 • 🛃	Page 🔹 🍈 Tools 👻 🂙
World Wide Web		~
World Wide Web		
The WorldWideWeb (W3) is a wide-area hypermedia information retrieval initiative aiming to give universal access to a large universe of	documents.	
Everything there is online about W3 is linked directly or indirectly to this document, including an <u>executive summary</u> of the project, <u>Mailir</u> <u>Questions</u> .	<u>g lists</u> , <u>Policy</u> , November's <u>W3 news</u> , <u>Frequently As</u>	<u>sked</u>
What's out there?		
Pointers to the world's online information, <u>subjects</u> , <u>W3 servers</u> , etc. <u>Help</u>		
on the browser you are using		
A list of W3 project components and their current state. (e.g. <u>Line Mode</u> , X11 <u>Viola</u> , <u>NeXTStep</u> , <u>Servers</u> , <u>Tools</u> , <u>Mail robot</u> ,	Library)	
Technical		
Details of protocols, formats, program internals etc Bibliography		
Paper documentation on W3 and references. People		
A list of some people involved in the project.		
History A summary of the history of the project.		
How can I help?		
If you would like to support the web Getting code		
Getting the code by <u>anonymous FTP</u> , etc.		
http://www.w3.org/History/19921103-hypertext/hype	rtext/WWW/TheProjec	t.html
	😱 🤤 Internet	
🛃 Start 🔰 🙆 🏈 💁 🙆 👋 😜 Fire 🗁 DIT 🔯 Micr 🖾 Aan 🌈 The 🌈 Http 🔪 Ado E		C 11:07 AM

WWW Project History - Windows Internet Explorer	
Ittp://www.w3.org/History/19921103-hypertext/hypertext/WWW/History.html	world first web page
🔁 👻	
Search the Internet - type your phrase here 🔽 🚯 Enter 🔹 🗗 Block PopUp 🥒 🕸 🐧 Shop Ebay 🗿 Learn How to Play Poker 🔹 튐	· · · ·
oogle 💽 world first web page 🛛 🕤 Go 🖗 🧔 🦃 🧭 🛜 🖈 👷 Bookmarks 🗸 🧟 2 blocked 🛛 🍄 Check 👻 🔨 AutoLink 👻 🔚 AutoFill 🍙 Send to 🗸	🗸 🥖 🔄 world 🔄 first 🔄 web 🔄 page 🛛 🔾 Sett
🛠 🏀 WWW Project History	🏠 🔻 🔝 👘 🖶 Page 👻 💮 Tools
History to date	
A few steps to date in the WWW project history are as follows:	
March 1989 First project proposal written and circulated for comment (TBL) . Paper "HyperText and CERN" (in ASCII or WriteNow form	at) produced as background.
October 1990	
Project proposal reformulated with encouragement from CN and ECP divisional management. <u>RC</u> is co-author.	
November 1990 Initial WorldWideWeb prototype developed on the NeXT <u>(TBL)</u> .	
November 1990	
Nicola Pellow joins and starts work on the line-mode browser . Bernd Pollermann helps get interface to CERNVM "FIND" inde	lex running. TBL gives a <u>colloquium</u> on hypertext in general.
'hristmas 1990	
Line mode and NeXTStep browsers demonstrable. Acces is possible to hypertext files, CERNVM "FIND", and internet news a	articles.
Febraury 1991	
workplan for the purposes of ECP division. 6 February 1991	
Presentation of the project to the ECP group.	
March 1991	
Line mode browser (www) released to limited audience on priam vax, rs6000, sun4.	
fay 1991	
Workplan produced for CN/AS group	
7 May 1991	
Presentation to C5 committee. General release of www on central CERN machines.	
2 June 1991	
CERN <u>Computer Seminar</u> on WWW. ugust 1991	
Files available on the net, posted on alt.hypertext (6, 16, 19th Aug), comp.sys.next (20th), comp.text.sgml and comp.mail.multi-	media (22nd) Jean-François Groff joins the project
Pies available on the net, posted on althypertext (0, 10, 19th Pag), compisystnext (20th), compitext sgin and compinalization.	filedia (2211d). <u>Scal-Francois Grou</u> Joins die project.
VMS/HELP and WAIS gateways installed. Mailing lists www-interest (now www-announce) and www-talk@info.cern.ch start	ted. One year status report. Anonymous telnet service started
December 1991	
Presented poster and demonstration at HT91 . W3 browser installed on VM/CMS. CERN computer newsletter announces W3	B to the HEP world.
5 January 1992	
Line mode browser release 1.1 available by anonymous FTP. See news . Presentation to AIHEP'92 at La Londe.	
2 February 1992	
Line mode v 1.2 annouced on alt.hypertext, comp.infosystems, comp.mail.multi-media, cern.sting, comp.archives.admin, and ma	ailing lists.
May 1992	
Presentation and <u>demo</u> at <u>JENC3</u> (Innsbruck). <u>Carl Barker</u> joins the project.	
June 1992	
ie de la constance de la const	🥡 🍚 Internet 🔍 100%

http://www.w3.org/MarkUp/tims_editor - Windows Internet Explorer	
C C + E http://www.w3.org/MarkUp/tims_editor	V 4 X tim bernes-lee
File Edit View Favorites Tools Help	
😭 🏟 🌈 http://www.w3.org/MarkUp/tims_editor	🟠 🔻 🔝 👻 🖶 Page 👻 🎯 Tools 👻 🎽
WorldWideWeb Into	Navigation
Info Precedence to of Version 0.13 Navigate Precedence to of Version 0.13 Document Precedence to of Version 0.13 Edit Precedence to of Version 0.13 Links Precedence to of Version 0.13 Document Precedence to of Version 0.13 An excercise in global information availability By Tim Bornors Location	<pre>< Previous Eack up Nc×t></pre>
Style Copyright 1950;91, CENN. Distribution restricted: ask for terms. TEOT VENDION CNLY Print P Page layout HyperText: Text which is not constrained to be linear. Windows P Services P	JFG's home page at InfoDesign
Services This version of the WWW application can pick up hyperlext information from files in a number of formals, from local files, from remote files using NF 5 or anonymous FTP, from hyperlext servers by name or keyword search and from internet news. Hype text has may be edited, and links made from hyperlext files to other files or any other information Style Image: Service of the file o	This home page demonstrates some simple concepts of the World-Wide Web <u>infrestructure</u> for global information sharing. Today, I showed Tim's original Web browser to <u>Marc Weber</u> . I just took some notes about him and I nked them to his name on the fly. I use this World-Wide Web editor exactly like any word processor, with the ocwer to create links from sensitive pieces of text to other places, for instance to personal notes or to any information source in the world.
Apply style 2 If you have any comments or have bugs, please mail timt/@infnicernichiquoting the varsion number (atrove) Bit is the editor HolWired: Whol's the neighborho wired 24 Ministration of the varsion number (atrove)	Geograph callict of <u>WWW servers</u> (at same list. Example : <u>CERN phone book</u> L Link to file WWW project documentation (at CERN Link to New N
style of selection >> in On the Right Apply style to selection Style sheet Open Flux Ned hears rumors of a deal between GN Apply style to all similar toot Save as Find unstyled text Net Suff Pro	A computer consultant, small publisher and freelance journalist who is currently investigating the early history of the Web, to be published in
Format Jerny Holzer's truisms take a turn for the We First line indert: 100 Successive indert: 130 Funt: Fe vetica Point Size: 12 GGM_ tag. Corporate recruiting form letters = t 130 Alt.cyn Cism What's the Point?	Wired. He loves HotWired. Welcome to HotWired!
Tabs:	
Done	🕡 🤤 Internet 🔍 100% 🔻 🖉
🛃 Start 🔰 🖉 5 Internet Explorer 🗸 🖻 🔯 🔯 Adobe Photoshop	o Alb 🛅 Ix\DITCHE 🛛 🐻 Microsoft PowerPoint 🤇 🖧 6:02 PM





Internet Archive, Library of Alexandria (copy of the Internet) mirror site

The Internet Archive

The Internet Archive at the Sibliothesa Alexandrina (BA) was established in April 2002 in cooperation with the Internet Archive in San Francisco. It is the second center worldwide, and exclusively the first outside the USA. The Petabox system, a custom designed system to store and manage one petabyle of data, is currently holding:

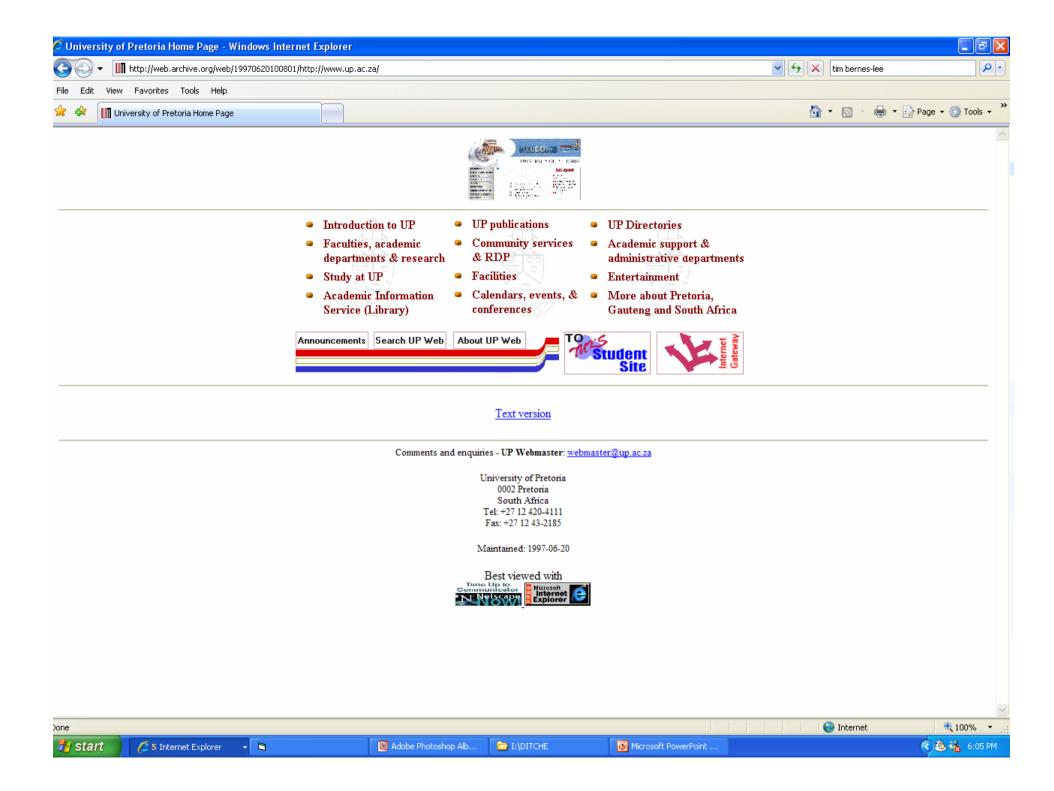
World Wide Web (WWW) snapshots from 1995 Over 50 billion pages

50.000 digitized books e existent's current capacity is 4.5 petabol 5 × 10⁶ bytes = 4.500.000 perceptions that

The data is stored on 29 rache are mounted on 29 rache

A rack can store up to: 100 million books in look of 12 million books in book format 12 million books in book format 12 million books in book format 100 archived films 1000 archived films 1000 archived films 1000 archived films





Internet	Archive Waybac	k Machine - Wind	dows Internet Exp	lorer							∎₽
• •	http://web.ar	chive.org/web/*/http:	://www.up.ac.za						🖌 🛃 🗙	tim bernes-lee	<u>م</u>
e Edit	View Favorites	Tools Help									
r 🕸 📗	Internet Archive V	Vayback Machine							6	• 🗟 - 🖶 • 🛙	🎐 Page 👻 🎯 Tools 👻
	ck Machine										
			Enter Web Addre	ess: http://		All 🔽 🔤	e Me Back	<u>iv. Search</u> <u>Compare A</u>	Archive Pages		
Searched	for <u>http://www</u>	/.up.ac.za									354 Results
denotes v	duplicates are no when site was up oically becomes a	dated.	onths after collectic		Results for	Jan 01, 1996	i - Nov 20, 20	007			
1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
0 pages	1 pages	3 pages	8 pages	12 pages	16 pages	21 pages	25 pages	60 pages	143 pages	4 pages	1 pages
	<u>Jun 20. 1997</u> *	<u>Jan 27, 1998</u> * <u>Dec 02, 1998</u> * <u>Dec 12, 1998</u> *	Jan 25, 1999 * Jan 25, 1999 * Jan 29, 1999 * Feb 08, 1999 * Apr 23, 1999 * Apr 23, 1999 * Oct 02, 1999 *	Mar 04, 2000 * May 10, 2000 * May 11, 2000 Jun 14, 2000 * Jun 19, 2000 Jun 21, 2000 Aug 15, 2000 Oct 06, 2000 * Oct 18, 2000 Nov 10, 2000 Dec 03, 2000 *	Jan 18, 2001 Feb 01, 2001 Feb 02, 2001 Feb 24, 2001 Feb 26, 2001 Mar 01, 2001 Mar 02, 2001 Mar 06, 2001 * Mar 31, 2001 May 16, 2001 May 16, 2001 Sep 05, 2001 * Nov 26, 2001 *	Feb 02, 2002 Apr 02, 2002 * May 25, 2002 * Jun 01, 2002 Aug 02, 2002 * Aug 07, 2002 Aug 13, 2002 Aug 20, 2002 Aug 20, 2002 Aug 20, 2002 Aug 27, 2002 Sep 28, 2002 * Sep 30, 2002 * Oct 19, 2002 Oct 25, 2002 * Nov 22, 2002 * Nov 22, 2002 * Nov 29, 2002 * Dec 26, 2002 *	Jan 24, 2003 Jan 26, 2003 * Mar 02, 2003 * Mar 19, 2003 * Apr 07, 2003 Apr 22, 2003 * Apr 23, 2003 May 28, 2003 Jun 08, 2003 * Jun 10, 2003 * Jun 10, 2003 * Jun 13, 2003 * Jun 18, 2003 * Jun 25, 2003 * Jul 25, 2003 * Aug 03, 2003 Aug 07, 2003 Oct 03, 2003 * Oct 10, 2003 * Oct 13, 2003 * Dec 14, 2003 * Dec 27, 2003	Feb 20, 2004 Mar 26, 2004 Apr 17, 2004 Jun 03, 2004 Jun 03, 2004 Jun 03, 2004 Jun 03, 2004 Jun 09, 2004 Jun 11, 2004 Jun 16, 2004 Jun 12, 2004 Jun 12, 2004 Jun 21, 2004 Jun 22, 2004 Jun 20, 2004 Jun 22, 2004 Jun 30, 2004 Jul 07, 2004 Jul 08, 2004 Jul 15, 2004 Jul 16, 2004 Jul 16, 2004 Jul 20, 2004 Jul 20, 2004 Jul 20, 2004 Jul 23, 2004 Jul 23, 2004 Jul 27, 2004 Aug 23, 2004 Aug 20, 2004	Jan 02, 2005 Jan 13, 2005 Jan 14, 2005 Jan 14, 2005 Jan 14, 2005 Jan 19, 2005 Jan 20, 2005 Jan 20, 2005 Jan 20, 2005 Jan 22, 2005 Jan 22, 2005 Jan 23, 2005 Feb 03, 2005 Feb 04, 2005 Feb 04, 2005 Feb 04, 2005 Feb 10, 2005 Feb 10, 2005 Feb 12, 2005 Feb 12, 2005 Feb 12, 2005 Feb 12, 2005 Feb 12, 2005 Feb 24, 2005 Feb 24, 2005 Feb 24, 2005 Feb 24, 2005 Mar 01, 2005 Mar 03, 2005 Mar 00, 2005 Mar 10, 2005 Mar 13, 2005 Mar 13, 2005 Mar 13, 2005	Mar 15, 2006 * Sep 23, 2006 * Nov 18, 2006 * Dec 08, 2006	<u>Jun 09. 2007</u> *
					Ac	lobe Photoshop Album	Starter Edition 3.2 - Ph	oto Downloader		🕘 Internet	🔍 100% ·
		net Explorer 🛛 🗸			Photoshop Alb	DITCHE		t PowerPoint			< 🗟 🐔 6:06 I

Use of digital information





- It took only 5 years for 50 million people to use the internet compared to 25 years for 50 million people to use phones
- Preservation or permanent availability of digital information is one of the processes affected by the evolution towards an all digital world



Definition of digital preservation



Digital preservation is a broad term used to describe the continued accessibility and maintenance of a digital resource safeguarding it into the foreseeable and the distant future. Digital preservation is a vital part of the creation and management of any digital collection.

http://www.tasi.ac.uk/advice/delivering/digpres.html



Joost sê sy handtekening is 'gecopy en gepaste'

Nellie Brand, Kaapstad

Joost van der Westhuizen se handtekening verskyn onderaan 'n brief van Progressive Investment Holdings (PIH) wat na bewering verkeerde inligting en 'n logo van 'n ander batebestuurder bevat en aan kliënte gestuur is.

PIH en drie verwante ondernemings is vandeesweek onder voorlopige kuratorskap geplaas weens onder meer die beweerde oortreding van die Wet op Finansiële Adviesen Tussengangerdienste (Fais).

Die Raad op Finansiële Dienste (RFD) het die aansoek om kuratorskap ingedien ná 'n jaarlange ondersoek na PIH se sake. Finansiële dienste is na bewering onregmatig gelewer.

Van der Westhuizen is sedert begin Maart bedryfshoof by PIH.

By navraag het hy aan *Sake24* gesê dat dit wel sy handtekening is wat onderaan die brief verskyn, maar bygevoeg dit is 'n afskrif daarvan wat op die brief "gecopy en gepaste" is.

Volgens hom het hy toestemming gegee dat sy handtekening gebruik mag word solank hy ingelig word waarvoor dit gebruik word.

Oor die betrokke brief, gedateer 5 Mei, het Van der Westhuizen gesê die bemarkingsdepartement van PIH het dit uitgestuur om aan mense te wys hy is "aan boord".

Volgens hom het hy nie geweet dit word "namens hom uitgestuur nie".

Van der Westhuisen het gesê hy weet nie wat die inhoud van die brief beteken nie. Hy het volgehou dat hy nog opgelei word en net 'n "student" is en niks van die onderneming se sake weet nie.

Hy ontvang opleiding van 'n persoonlike opleier, mnr. Murray Kilgour, sodat hy in Januarie volgende jaar as uitvoerende hoof van die maatskappy kan oorneem.

Die pos van bedryfshoof is intussen vir hom geskep deur mnr. Freddie Andalaft, huidige uitvoerende markingsdoeleindes, nie vir kontrakte nie," het hy gesê.

Volgens hom word dit ook op pamflette en penne gebruik, asook bemarkingsbriewe soos die betrokke een.

Dié brief bevat ook 'n handelsmerk van die batebestuurder Dynamic Wealth wat volgens mnr. Phillip Hattingh, besturende direkteur van die onderneming, onwettig gebruik is.

Hattingh het gesê hulle gaan regstappe doen daaroor. Volgens hom word alle briewe waarop hul handelsmerk verskyn, deur hul nakomingskantoor goedgekeur vanweë die streng vereistes wat deur die Fais-wet gestel word vir inligting wat aan kliënte gestuur word.

Dié brief se inligting is ook verkeerd, volgens Hattingh.

In die brief word onder meer gesê dat van Dynamic Wealth se "aandeleplatform" gebruik gemaak word.

Volgens Hattingh het PIH nog nooit 'n aandelerekening by Dynamic Wealth gehad nie.

Andalaft het gesê dié logo is gebruik na aanleiding van 'n brief van 'n direkteur van Dynamic Wealth.

In daardie brief, gedateer April, skryf mnr. Kobus Zietsman, 'n direkteur van Dynamic Wealth, dat hulle die effektetrusts van Progressive Asset Management (PAM) so spoedig moontlik gaan oorneem.

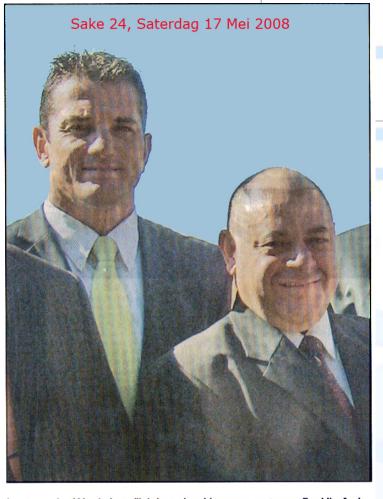
Hattingh het egter gesê niks is deur die direksie goedgekeur nie.

Hattingh het gesê dit is nie die eerste keer dat sulke ondernemings probeer om uit Dynamic Wealth se goeie naam munt te slaan nie.

Dit het ook al met ander bekender batebestuurders gebeur, het hy gesê.

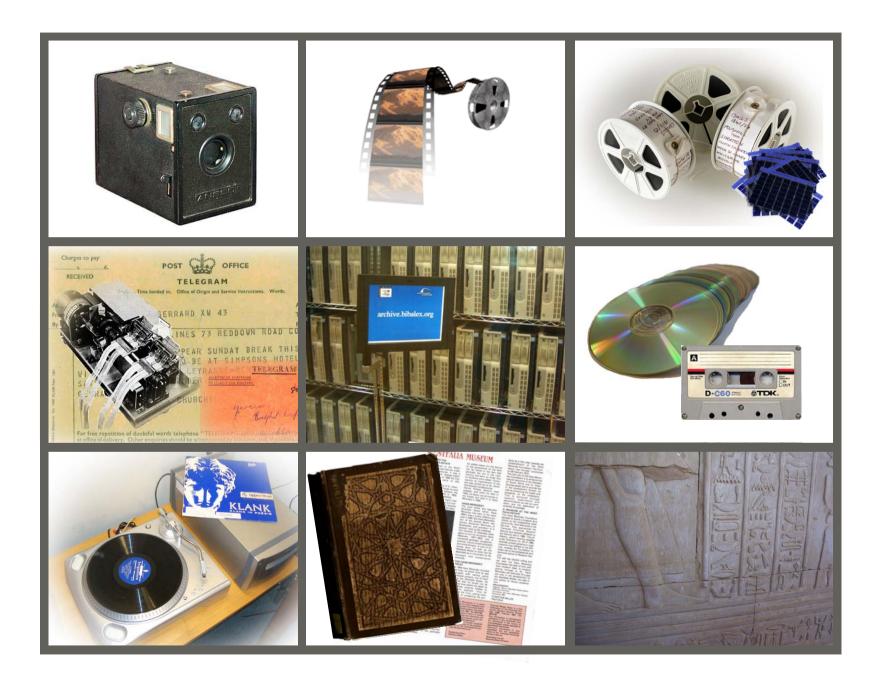
Die RFD het intussen 'n verklaring uitgereik waarin gesê is die betrokke effektetrusts word deur Metropolitan Collective Investments besit.

Die RFD het gesê dié effektetrusts is nie onder kuratorskap nie. Dit is



Joost van der Westhuizen (links) verskyn hier saam met mnr. Freddie Andalaft op 'n foto wat onlangs by 'n ontbytgeselligheid in Bloemfontein geneem is. Foto: VIDA BOOYSEN





Future of academic libraries

No.1 assumption (ACRL, March 2007)

There will be an

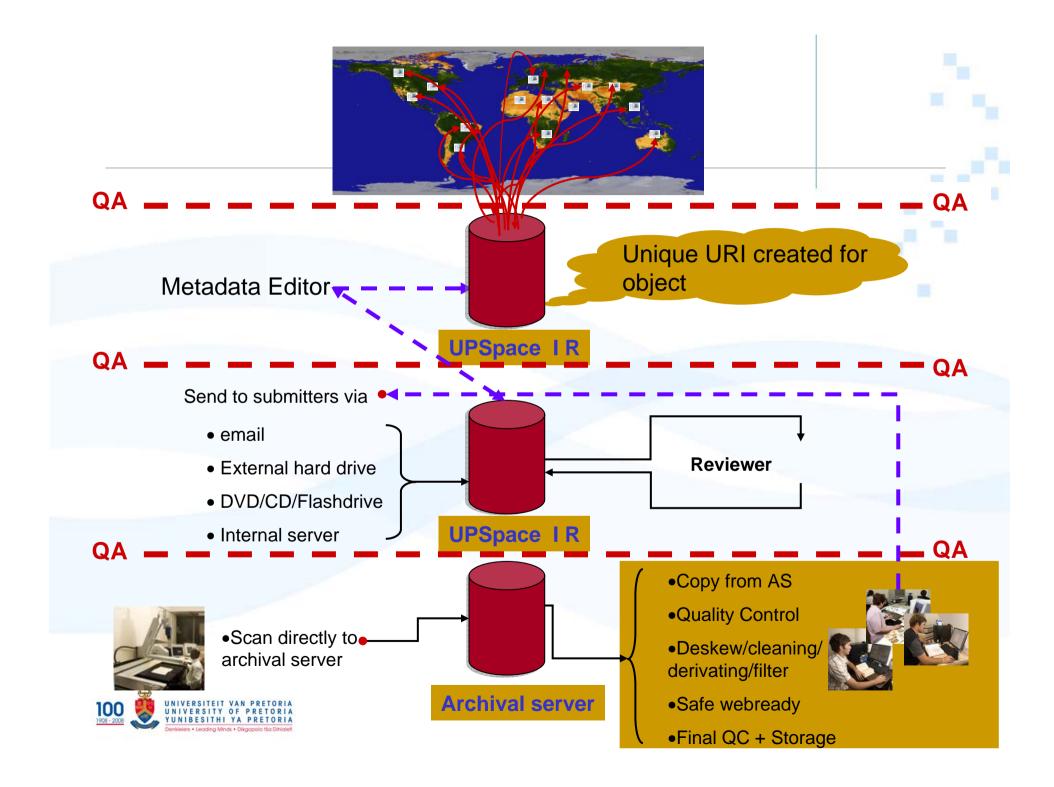
- preserving digital archives, and
- increased emphasis on digitizing collections
- improving methods of data storage and retrieval
- The digitization of unique print collections may emerge as one of the primary missions of academic libraries in the 21st century
- Librarians should collaborate with disciplinary colleagues in the curation of data as part of the research process

http://www.ala.org/ala/acrl/acrlpubs/crlnews/backissues2007/april07/tenassumptions.cfm



Digital workflow of the Alexandria Library. Software for this workflow is available at http://wiki.bibalex.org/DAFWiki/index.php/Main_Page





Standards (1)

- Preservation Metadata Framework Working Group (OCLC, 2003)
- PREMIS (2005)
- OAIS (Open Archival Information System)
- Z39.87 Standard for Technical Metadata for Digital Still Images(ANSI/NISO)



Standards (2)

- ISO (International Standards Organisation) defines "quality as the totality of characteristics of an entity that bear on its ability to satisfy stated or implied needs" (ISO8042:1994)
- or "the degree to which a set of inherent characteristics fullfils requirements" (ISO9000:2000)



Preservation Metadata Framework Working Group (Report 2003)

Framework for research

- Outline the types of information that should be associated with an archived digital object
- The use of metadata to support the digital preservation process

http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf



PREMIS Working Group (2005)



The PREMIS (Preservation Metadata: Implementation Strategies Working Group -

- Develop a data dictionary of core elements for archived objects
- Guide the implementation of element sets in preservation systems
- Suggest best practice for populating the elements

http://www.oclc.org/research/projects/pmwg/pm_framework.pdf



OAIS (Open Archival Information System)



- The OAIS (Open Archival Information System) reference model was developed under the auspices of NASA's Consultative Committee for Space Data Systems (CCSDS)
- The OAIS reference model is a conceptual framework for a digital archive
- Regarded as the "standard" for digital object repositories



Z39.87 - Standard for Technical Metadata for Digital Still Images (NISO & AIIM)

Z39.87 is a standard which defines a set of metadata elements for raster digital images

The purpose is to help in the development, exchange and interpretation of digital images

The original DIG35 goals were adapted by the NISO group





Scanning

- No set resolution can be selected for all projects
- Resolution for a master image range between 300 -600 dpi
- Colour settings 8-bit greyscale; 24-bit colour
- The most widely adopted format for storing a preservation quality digital master is uncompressed TIFF



Derivative image

- A derivative is a manipulated image derived from the master image, to produce smaller file sizes
- Lossy file formats such as JPEG are used for derivative images
- Resolution ranges between 72 dpi and 150 dpi and up to 800 pixels in width
- ICC (International Colour Consortium) profiles



Reasons for preservation

- Updated versions of the file format
- Reading device become obsolute
- Updated versions of the software used to create, manage, or access digital content
- Changes in computers
- Movement at vendors level
- Unforeseen errors



Requirements of data protection

- Visibility/accessibility
- Regular quality control
- Authenticity
- Security
- Performance
- Ease of use
- Interoperability
- Cost of ownership
- Automation

Web Buyers Guide, 31-03-08



Preservation methods



Refreshing:

Copy the same type of digital information from one long-term storage medium to another

- Modified refreshing: Copy information to another medium of a similar type
- Refreshing is part of a process or program
- Refreshing address issues such as decay and obsolescence



Migration and Emulation



Move or adapt the objects to another platform

• Emulating:

Environment will be adapted to new platform (the objects themselves will not be tampered with)



Preserve the usability of a .TIFF file



- A TIFF viewer, plus its formal specification and sufficient subsidiary documentation to explain how it work in practice must be preserved
- To run the TIFF viewer an operating system must be preserved
- To run the operating system -
 - the original hardware will need to be preserved, or
 - emulation software that allows the old hardware to be emulated on new machines needs to be developed



PRONOM Welcome - Windows	Internet Explorer		
💽 🗸 🛓 http://www.nationalarchives.gov.	uk/pronom/	🗸 🛃 🔀 🖌	oogle 🛛 🔎 🗸
Image: State of the state		· 🔐 • 🔊 •	🖶 🔹 🔂 Page 🔻 🎯 Tools 👻 🔞 🗸
Contact us Help The National Archives	> A to Z index > Site search > Accessib	ility	<u>MyPage Sign in Register</u>
Home A u	bout Visit Research, education & Search the s s us online exhibitions archives p	Services for News Shop Professionals online	
You are here: <u>Home</u> > <u>Services for professiona</u>	s > Preservation > PRONOM		
The technic	RONOM		bout 🕀 Add an entry Help 🔝 Information resources
	PRONOM is a resource for anyone requiring imp onents required to support long-term access to		
Tools and Services Free PRONOM tools and services to support digital preservation, including DROID, the automatic file format identification tool, together with links to relevant external tools and services. Tools and Services.	Contribute to PRONOM Contribute new information to PRONOM via our online <u>submission form</u> > Find out more about <u>PRONOM's creators</u> >	 New to PRONOM? What is PRONOM? How do I search PRONOM Who is PRONOM for? How do I find out more? 	
Terms of use Copy The National Ar	right Privacy Complaints procedure Fre	edom of Information Top of -44 (0) 20 8876 3444. Contact u	page ^ s
Done		l 😜 Intern	et 🔍 100% 🝷 🧋
🛃 start 🔄 KARM 💽 Microsof	🚺 Vodafon 🌈 PRONOM 🧷 Des	ktop 🤌 📑 🖸 🖬 🛒 📕 🧐 🍇 🌱	🕻 🗞 🏧 🏷 🐵 😓 📨 繘 😵 01:07 PM -

PRONOM Simple Search - Windows Internet Explorer		
🚱 🕤 🔹 http://www.nationalarchives.gov.uk/PRONOM/BasicSearch/proBasicSearch.aspx?status=new	🖌 🗲 🗙 Google	₽ -
See		
🛠 🅸 🗚 PRONOM Simple Search	🟠 • 🔊 - 🖶	🔹 📴 Page 👻 🍈 Tools 👻 🔞 🗸
Contact us Help A to Z index Site search Accessibility		~
The National Archives	MyE	Page <u>Sign in</u> <u>Register</u>
Home About Visit Research, education & Search the Services for us us online exhibitions archives professionals	News Shop online	
You are here: <u>Home</u> > <u>Services for professionals</u> > <u>Preservation</u> > <u>PRONOM</u> > Search: PRONOM		
The technical registry PRONOM		 Add an entry Information resources
Search : Simple Search		?Help : Simple Search
Simple search File format PRONOM Unique Identifier Software Vendor Lifecycles		
1. Search		
Enter a simple search string and then click 'search'JPEG Search		
Terms of use Copyright Privacy Complaints procedure Freedom of Inf The National Archives, Kew, Richmond, Surrey, TW9 4DU. Tel: +44 (0) 20 88		^
Rookmark with Delicious 🖓 Digg 📑 Eacoback 📿 Coogle 👼 roddit 🔊	Stumblelinen 32 Vaboa	
Bookmark with: 📲 <u>Delicious</u> 😭 <u>Digg</u> 🖪 <u>Facebook</u> <u>G</u> <u>Google</u> 🧐 <u>reddit</u> 👽 (what are these links?)		
		~
Done	🌍 Internet	🔍 100% 🔻 💡
🛃 start 🔹 🖄 KARM 🛛 💽 Microsof 💽 Vodafon 💋 PRONOM 🧭 Desktop 🎽 🚅	0 • • • • • • • • • • • • • • • • • • •	😵 😌 😓 🔤 🗖 😵 - 01:10 PM

PRONOM Simple Search - Windows Internet Explorer	- 8 🗙
🚱 🕤 👻 http://www.nationalarchives.gov.uk/PRONOM/BasicSearch/proBasicSearch.aspx?status=listReport 🛛 🔽 🖌 Google	P -
😪 < File Edit View Favorites Tools Help	
🛠 🏟 🗚 PRONOM Simple Search	r 🔘 Tools 🔻 💽 🕇
Simple search File format PRONOM Unique Identifier Software Vendor Lifecycles	^
page 1 <u>2</u> () ()	Print
JPEG File Interchange Format (1.00)	
The JPEG File Interchange Format (JFIF) is a file format for storing JPEG-compressed raster images. It was developed by the Independent JPEG Group and C-Cube Microsystems, in the absence of any such format being defined in the JPEG standard, and r became a de facto standard; this is what is commonly referred to as the JPEG file format. A JFIF file comprises a JPEG data streat together with a JFIF marker. It begins with a Start of Image (SOI) marker, immediately followed by a JFIF Application (APPO). Th followed by the JPEG image data, which is terminated by an End of Image (EOI) marker. JFIF supports up to 24-bit colour and us lossy compression (based on the Discrete Cosine Transform algorithm). Other types of compression are available through JPEG extensions, including progressive image buildup, arithmetic encoding, variable quantization, selective refinement, image tiling, and lossless compression, but these may not be supported by all JFIF readers and writers.	am iis is ies
JPEG File Interchange Format (1.01)	
The JPEG File Interchange Format (JFIF) is a file format for storing JPEG-compressed raster images. It was developed by the Independent JPEG Group and C-Cube Microsystems, in the absence of any such format being defined in the JPEG standard, and rebecame a de facto standard; this is what is commonly referred to as the JPEG file format. A JFIF file comprises a JPEG data streat together with a JFIF marker. It begins with a Start of Image (SOI) marker, immediately followed by a JFIF Application (APP0). The followed by the JPEG image data, which is terminated by an End of Image (EOI) marker. JFIF supports up to 24-bit colour and us lossy compression (based on the Discrete Cosine Transform algorithm). Other types of compression are available through JPEG extensions, including progressive image buildup, arithmetic encoding, variable quantization, selective refinement, image tiling, and lossless compression, but these may not be supported by all JFIF readers and writers.	am iis is ies
JPEG File Interchange Format (1.02)	
The JPEG File Interchange Format (JFIF) is a file format for storing JPEG-compressed raster images. It was developed by the Independent JPEG Group and C-Cube Microsystems, in the absence of any such format being defined in the JPEG standard, and r became a de facto standard. A JFIF file comprises a JPEG data stream together with a JFIF marker. It begins with a Start of Ima (SOI) marker, immediately followed by a JFIF Application (APPO) marker and one or more optional application extension markers. T followed by the JPEG image data, which is terminated by an End of Image (EOI) marker. JFIF supports up to 24-bit colour and us lossy compression (based on the Discrete Cosine Transform algorithm). Other types of compression are available through JPEG extensions, including progressive image buildup, arithmetic encoding, variable quantization, selective refinement, image tiling, and lossless compression, but these may not be supported by all JFIF readers and writers.	ige This is ises
See	🔍 100% 🔹 🛒
🛃 Start 🔰 🖎 KARM 🛛 🐼 Microsof 🚺 Vodafon 🧷 PRONOM 🦿 Desktop 🎽 🚅 🖳 🖬 🖉 🖬 🐼 🖓 🍇 🙀 🇞 🚳 😓 📼	🗖 🍄 01:11 PM

A PRONOM Search by form	at	🟠 🔹 🔝 🕤 🖶 🖬 Page 🕶 🎯 Too
Details for: JPX (JPEG 2		
· · · · · · · · · · · · · · · · · · ·)00 Extended)	🕒 Save as XML CSV 🛛 🚐 Print
Go to: Summary Docum > Properties >	entation 〉 Signatures 〉 Compression 〉 Cha	aracter encoding > Rights > Reference files
Summary		
Name Version	JPX (JPEG 2000 Extended)	
Other names	JPF	
Identifiers	PUID: fmt/151	
Family	POID. 111(151	
Classification		
Disclosure		
Description	format for the JPEG 2000 international standard Experts Group in 2000 for the compression of p standard specifies how an image and its metad is widely used for storing and transmitting phot it is based on JP2 it can support multiple layers	, lossless, extension of the JP2 format, the main file rd for image coding, created by the Joint Photographic photographic images for storage or transmission. The data is transformed into byte stream data. This format tos and other compressed image data online. Although s, animation and other features, and is a lossless made up of a contiguous sequence of boxes, beginning h provides version and file type information.
Orientation	Text	
Byte order	Big-endian (Motorola)	
byte order		

Preservation of the format



Digital formats contain texts, databases, still and moving images, audio, graphics, software and web pages. They are fragile and require purposeful production, maintenance and management to be retained

- Viability maintenance of the bitstream
- *Renderability* viewable by humans and processible by computers
- Understandability interpretable by humans

http://www.icpsr.umich.edu/dpm/dpm-eng/terminology/preservation.html

Part II: Preservation & Trusted Digital Repositories

Ina Smith



Institutional Repository

"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."

Clifford A. Lynch,

"Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age" ARL, no. 226 (February 2003): 1-7.



Digitally born & digitized material







CLINICAL IMAGES

Awaking a sleeping epidemic

Riaan van Coller, Elna van Reusburg , Clars Schutte, Deleue Brink, Gerhard Welthagen, M G Dove

Two patients with African sleeping sickness (S5) presented to the neurology unit, Protocia Academic Hospital, during 2004 and 2005. S5 has shown a recent resurgence, with epidemics in the Sudar, Angola and the Democratic Republic of Congo. The number of infected people in Africa is currently estimated at more than 500 000. According to the World Health Organization (WHO), about 20 Trypurasense bracei previous and 300 T. 5. rholzenias infections are diagnosed yearly outside endemic areas in Africa. Migration, tourism, peacekeeping and military interventions and the re-emergence of S5 epidemics might increase these numbers.³

The electroencephalogram (EBG) is often useful in the diagnosis of coma and delirium, but has not been widely used in the diagnosis of SS. The EEG is proposed as a novel way to follow disease progression, treatment response and treatmentinduced encephalography.

Case 1

250

A 27-year-old man presented with a 4-month history of fatigue, loss of appetite, intermittent severe handaches, excessive daytime sleepiness, loss of concentration and insomnis. He had traveled to Malswi 8 months before admission. His temperature was 38.9°C, he had a palpable hepatomegaly and an unremarkable neurological examination although his cognitive response was slow. Disgnosis of African typencsomiasi was made on a Gienna-stained blood smeet (Fig. 1).

Shortly after admission the patient had a tonic-donic seizure, with peak-ictal confusion. Treatment with surnaria was started and repeat blood smears a that 48 hours were clear of trypanosome. The cerebrospinal fluid (CSF) showed no trypanosemes but a total protein level of 1.2 g/l, glucose 2.1 mmol/l, 4 polymorphs and 82 hymphocytes. WHD-recommended treatment with malanceprol was started.¹

Since no trypanosomes were isolated from inoculated mice the diagnosis of West African trypanosomiasis (WAT) was made. Efforithing was unavailable and treatment with melanoprol continued. The patient recovered well and returned to the UK.

The authors are in the departments of Neurology and Clinical Microbiology, University of Pretoria. Five months after discharge he presented to the Hospital for Tropical Diseases in London with fever, alsopiness and an active CSF Diagnosis of a relapse was made which presed a diagnostic dilemma – recurring T. b. genbinue. Treatment with effornibles was given which cleared his condition.

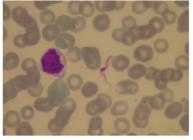


Fig. 1. Gimme sour (case 1) showing extracellular trypmononissis peresite in the peripherel blood.

Case 2

A 53-year-old man presented with a 2-weak history of fever, headsche and epicodic contusion. He was a farmer from Nariba in Zimbelwe where he had been treated for malaria without any dinical improvement. He gave a history of multiple betwe By bites but did not have a chance. A Gamas-stained blood smear showed Trypurcesson spp. On admission his temperature was 39 AFC, but the general examination was unremarkable. He was very sleepy but easily arounable. The diagnosis of East African trypunceanises (EAT) was confirmed by isolating T. Ardeoiswes from inoculated mice. Treatment with surramin was started and repeated Gienna-stained blood smears did not show any trypanosume.

Our first patient had a series of EBG recordings. These indicated a low-voltage mixed-frequency background with episodic, generalised but frontally dominant irregular delta activity (Fig. 2). Follow-up showed a gradual improvement in the frequency of the background. The second existin had an EBG



_ 8 ×

1

Digitally born & digitized material

S. Afr. J. Agric. Ext., 1993 : 47-54 S. Afr. Tvdskr. Landbouvoori., 1993 : 47-54

THE ROLE OF EXTENSION IN TRADITIONAL AGRICULTURE : EVIDENCE FROM THE FARMER SUPPORT PROGRAMME

J F Kirsten², J van Zyi³ and HJ Sartorius von Bach⁴

ABSTRACT

This paper evaluates the extension and training element of the FSP as implemented in the farmer support programmes of Venda and Lebowa. The importance and contribution of extension and training in the success of the programme can be judged from this paper. The general conclusion of the paper is that extension and training play an important role in FSPs. They are closely associated with increased production. However, to what extent it contributed towards increased production remains a point of contention. Some analysts and observers argue that only the rural elite has access to FSP related services. Results from the analyses in the paper, however, show that comparable households achieve higher yields when they get appropriate extension. Training and extension thus at least partly contribute to higher maize yields in FSP areas.

UITTREKSEL

Hierdie artikel beskou die opleidings en voorligtingskomponent van die kleinboer ondersteuningsprogram ("Farmer Support Programme' - "FSP") soos dit in sekere gebiede van Venda en Lebowa geïmplimenteer is. Die belangrikheid en bydrae van voorligting en opleiding in die sukses van hierdie program biyk duidelik uit hierdie artikel. Die gevolgtrekking word gemaak dat voorligting en opleiding een van die belangrikste elemente van hierdie program is en nou geassosieer word met verhoging in produksie. Tot watter mate opleiding en voorligting tot die verhoging in produksie bydra, bly egter 'n punt van dispuut. Ontledings in die artikel toon egter aan dat vergelykbare huishoudings meer produseer Indien hulle gepaste voorligting ontvang. Dit wil dus voorkom asof opleiding en voorligting ten minste gedeeltelik bydra tot hoër opbrengste in die gebiede waar die kleinboer teuningsprogramme geïmplimenteer is.

1. INTRODUCTION

In response to the ineffective and costly large scale project approach in homeland agriculture, the Development Bank of Southern Africa introduced the concept of a small holder farmer support approach to aid the development of black agriculture in South Africa. Considering the various constraints faced by small farmers in the homelands (cf. Van Rooyen et al, 1987), the farmer support programme (FSP) was

by improving farmers' access to support services over a broad base in a sequential and evolutionary manner" (Van Rooyen, 1993).

In order to reach this objective the FSP comprises six basic elements, i.e. the supply of inputs and capital to farmers, mechanisation services, marketing services, extension services, training and

>gi|73624990|gb|DQ103533.1| Botryosphaeria rhodina strain CMW9074 185 ribosomal RNA gene GGAAGGATCATTACCGAGTTTTCGAGCTCCGGCTCGACCTCTCCCACCCTTTGTGAACGTACCTCTGTTGC AATAAACTAAAAACTTTCAACAACGGATCTCTTGGTTCTGGCATCGATGAAGAACGCAGCGAAATGCGATA AGTAATGTGAATTGCAGAATTCAGTGAATCATCGAATCTTTGAACGCACATTGCGCCCCTTGGTATTCCG gggggcatgcctgttcgagcgtcattacaaccctcaagctctgcttggaattgggcaccgtcctcactgc GGACGCGCCTCAAAGACCTCGGCGGTGGCTGTTCAGCCCTCAAGCGTAGAATACACCTCGCTTTGGA GCGGTTGGCGTCGCCCGCCGGACGAACCTTCTGAACTTTTCTCAAGGTTGACCTCGGAT



Ġ Back • 🐑 - 💌	💈 🏠 🔎 Search hrvorit	tes 🚱 🔗 🌺	🗔 🛄 📊 🕯	🛍 💽 💉	-8	
Address 👌 http://www.dspac	e.up.ac.za/samedia.txt				Go Links	» 🤹 •
Y! - @-	💌 Search Web 👻 🖉 🗖	🗟 - 🅀 - 🔞	+ 🚳 My Yahoo! 🔸	• 😡 Answers 🔹	de Games	- >>
Sent: 17 February : To: Ina Smith	er [elsabe.olivier8up.ac. 2006 11:49 pestemming van SAMedia vi		n Collection			
To: <elsabe.olivie< td=""><th>der Berg" «VDBergH.BIB§ma têup.ac.za» uary 17, 2006 10:48 AM</th><th></th><th></th><th></th><th></th><td></td></elsabe.olivie<>	der Berg" «VDBergH.BIB§ma têup.ac.za» uary 17, 2006 10:48 AM					

File Edit View Favorites Tools Help

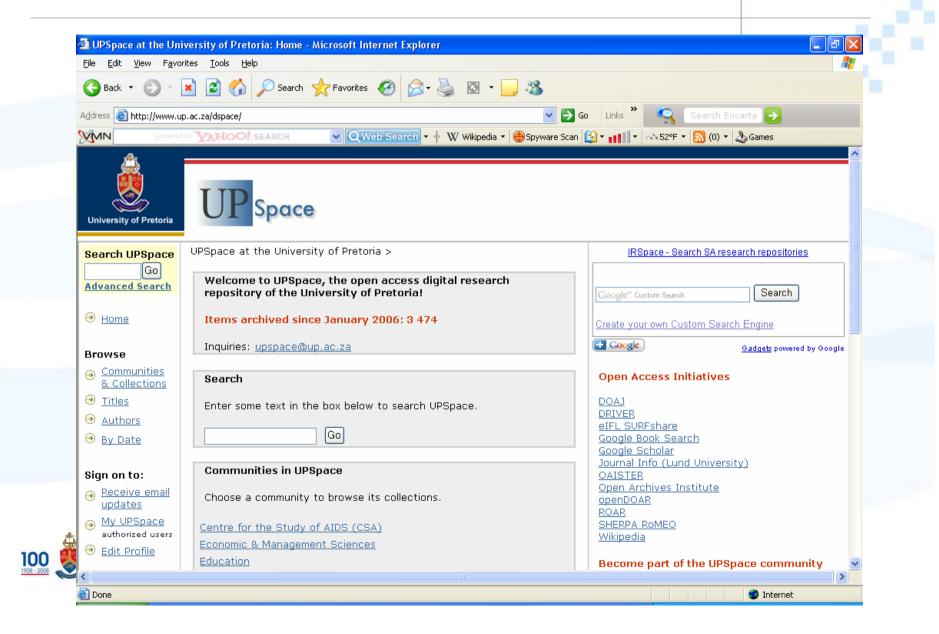
Elsabe In oorleg met my direkte hoof gee ons aan julle die vergunning om artikels uit S& Media se databasis op te neem in julle Jonathan Jansen Collection. > Hester van der Berg > Besuurder: Så Media 051 401 2305

University of the Free State: This message and its contents are subject to > a disclaimer. Local intranet

🕑 Windows Media Player



https://www.up.ac.za/dspace/



Benefits of an open access IR



- Research out quickly, worldwide
- Increases visibility, usage, impact of research "open access [material] are read more widely, and, therefore, cited more frequently. The consequence of this is that they have greater impact" (Jones, Andrew and MacColl 2006)
- Open access to all also those who cannot afford subscribing
- Persistent URL
- Decentralised/ Distributed input
- E-workflow for quality control
- Full text searchable
- Central archive of research
- Preservation function



11 Repositories National (ROAR)

- African Higher Education Research Online
- CSIR Research Space
- Durban University of Technology Institutional Repository
- Rhodes eResearch Repository
- Stellenbosch University Electronic Theses & Dissertations
- University of Cape Town Computer Science Research Document Archive
- University of Cape Town Lawspace
- University of Johannesburg Electronic Theses & Dissertations
- University of Pretoria Electronic Theses & Dissertations
- University of Pretoria Institutional Repository (UPSpace)
- University of the Western Cape Electronic Theses & Dissertations



1 200 Repositories Internationally



Univ. of Australia

Inth DBpace University of St Andrews Digital Research Repository > Welcome to the University of St Andrews Digital Research Store September 2006 the University regards these to be submitted to the metitive months content will give as an wead of the to these, research	Andrews. From
Welcome to the University of St Andrews Digital Reset This is a digital reporting of research output from the University of St September 2006 the University engages theses to be submitted to the	arch Repository electron thesis
September 2006 the University requires theses to be submitted to the	Andrews. From
	h articles and papers. Septembe
owse	2006 you are require
Communities Search	to submit
Enter some text in the box below to search DSpace.	copy of y thesis
Authors Go)	Follow the
By Data	links for b with your
n on to: Communities in DSpace	submission and more
ficent mill	informatio
updates	electronic and poot
Ma Alapiese Art History (School of) extensional users Biology (School of)	theses
Builds Aller at Without	How to
Help Chemistry (School of)	submit
	papers
Library Home Computer Science (School of)	Post of the



Cornell University

And many more

Trusted Repository Defined



"One whose mission is to provide **reliable**, **long-term access** to managed digital resources to its designated community, now and in the future."

(RLG-OCLC Report 2002)



Attributes of a Trusted Repository

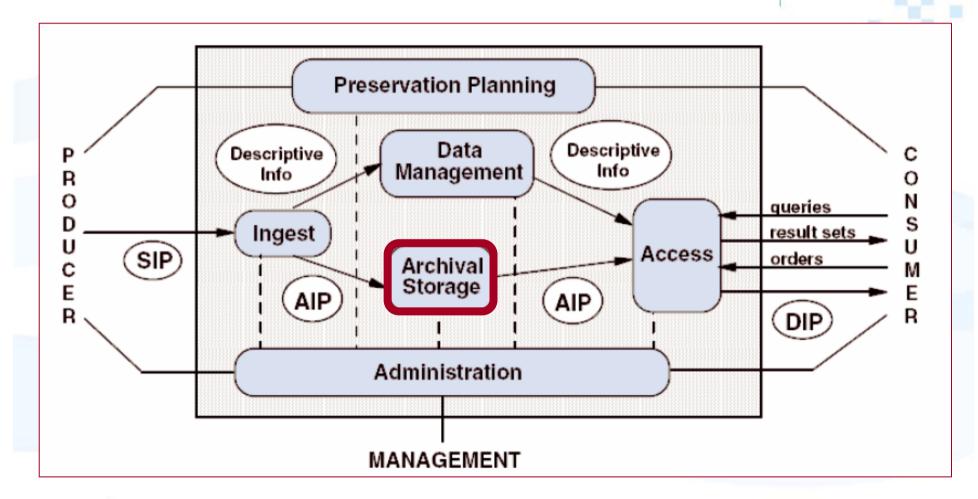
- Compliance with the Reference Model for an Open Archival Information System (OAIS)
- Administrative responsibility
- Organizational viability
- Financial sustainability
- Technological & procedural suitability
- System security
- Procedural accountability

Source: Trusted Digital Repositories: Attributes and Responsibilities An RLG-OCLC Report <u>http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf</u>





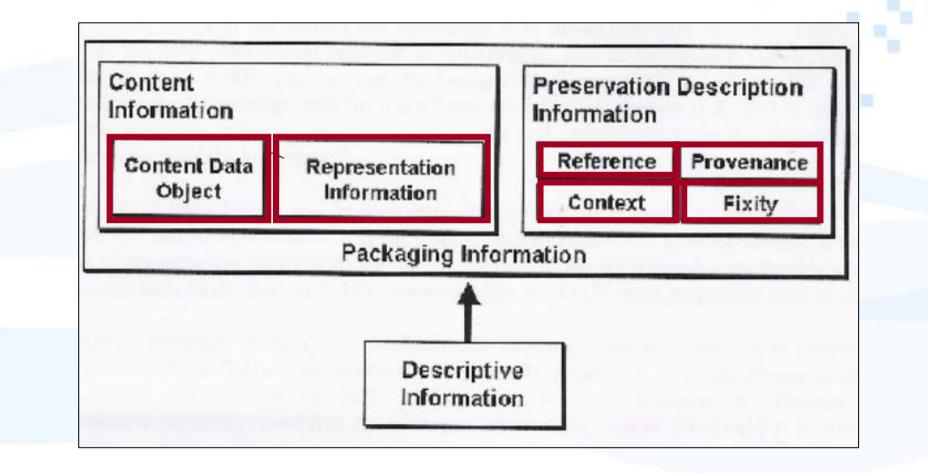
OAIS Functional Model – Archival Storage





Source: http://public.ccsds.org/publications/archive/650x0b1.pdf

Archival Information Package (Digital item submitted)





Source: http://public.ccsds.org/publications/archive/650x0b1.pdf



Technologies for enabling trust & preservation



Digital Repository Software

- Proquest Digital Commons (proprietary)
- DSpace (open source)
- Content^{DM} (proprietary)
- Fedora (open source)
- E-Prints (open source)
- Greenstone (open source)





File formats

- Proprietary e.g. MSWord
- Open formats e.g. ASCII plain text no restrictions
- Industry standard formats e.g. HTML, PDF
- Industry/ Open standard format with proprietary extension e.g. Microsoft Version of XML



DSpace Commitment to Preservation

- 2 levels of preservation: Bit & Functional
- Three levels of preservation for a given file format:
 - Supported: The format will be fully supported and preserved using either format migration or emulation techniques.
 - Known: The format can be recognised by DSpace, but full support cannot be guaranteed.
 - Unsupported: The format cannot be recognised by DSpace; these will be listed as "application/octet-stream", aka Unknown.
- Bit-level preservation will be done so that digital archaeologists of the future will have the raw material to work with if the material proves to be worth that effort.

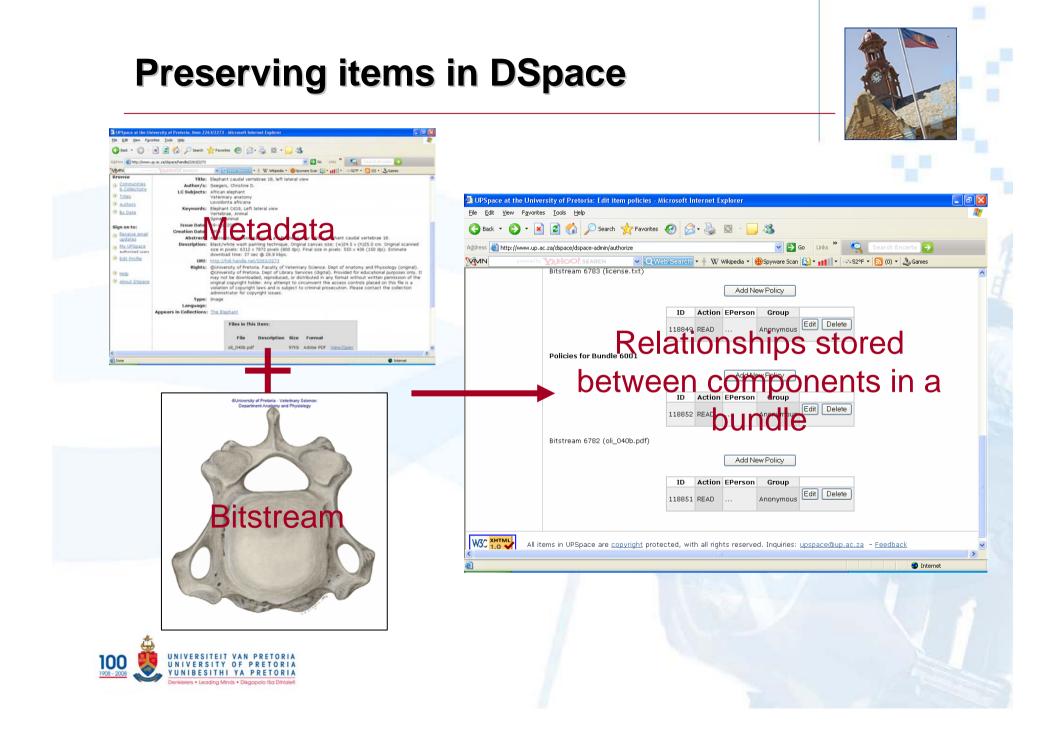


UPSpace at the University	rsity of Pretoria: Bitstream Format Registry - Microsoft Internet Explorer	PX
<u>File E</u> dit <u>V</u> iew F <u>a</u> vorites	s <u>T</u> ools <u>H</u> elp	-
🚱 Back 🝷 🛞 🕤 💌	📔 🟠 🔎 Search 🧙 Favorites 🚱 🔗 + 🌺 🔯 + 🔜 🖄	
Address 🙆 http://www.up.ad	ic.za/dspace/dspace-admin/format-registry 💽 🎦 🔽 🛛 Links 🎽 🔍 Search Encarta 🕞	
powered by	YAHOO! SEARCH V QWeb Search • 🔶 W Wikipedia • 🖶 Spyware Scan 🔯 • 📲 🔹 🗫 52°F • 🔊 (0) • & Games	
⊖ <u>Groups</u>	Bitstream Format Registry	~
 Items Dublin Core Registry 	Extensions are comma-separated list: E.g. Adobe PDF, XML, Text, HTML, ad files	5.
Bitstream Format Registry Workflow	When you add a bitstream format, it is	
Authorization Edit News	Help	
	ID / MIME Type / Name / Lor On / Support Level / Internal? / Extensions 1 application/octet-: Unknown Unknown da Unknown V	
ອ <u>Statistics</u>		
⊛ <u>Help</u>	2 text/plain License Item-s fic license ag Known 🗸 🔽 Update Delete	
⊖ <u>Log Out</u>	✓ 3 application/pdf Adobe PDF Adobe Portable Docum Known ▼ □ pdf Update Delete	
	4 text/xml XML Extensible Markup Lang Known 🕑 🗆 xml Update Delete	
	5 text/plain Text Plain Text Known 🕑 🗖 txt, asc Update Delete	
	6 text/html HTML Hypertext Markup Langi, Known 💌 🗆 htm, html Update Delete	
<	7 application/mswo Microsoft Worl Microsoft Word Known 🗸 🗖 doc Update Delete	× >
Done	🔮 Internet	

UPSpace Policy for file formats

- Everything put in UPSpace will be retrievable
- As many files formats as possible will be recognised
- As many known file formats as possible will be supported through UPSpace
- Formats and techniques will be continuously monitored to ensure needs can be accommodated as they arise
- The size of a bitstream allowed for submission is currently unlimited, but this will be revised over time
- The same file can be submitted in more than one format, of which one must be pdf (does not apply to media files)





Metadata

- Data about data
- Qualified Dublin Core Metadata Schema
- DSpace supports the Open Archives Initiative's Protocol for Metadata Harvesting (OAI-PMH) v2.0 as a data provider
- Enhance Descriptive Metadata
- Capture Administrative Metadata (incl. preservation metadata)

"Preservation metadata is the information necessary to maintain the viability, renderability, and understandibility of digital resources over the long-term."

Source: Feasibility and Requirements Study on Preservation of E-Prints/ Hamish et al.



AIP – Content Info – Representation Info

Preservation Metadata



identifier	uri	http://hdl.handle.net/2263/4726
description	abstract	In this paper, an attempt will be made examine the concept ataraxia as it appe the works of Pyrro of Elis, Sextus Empi
description	provenance	Scanned in 24-bit descreened colour 100% or Annual Scanner at 400 dpi.
description	provenance	Submitted by Ina Smith (ina@ais.up.ac.za) on 2008-03-13T11:59:46Z No. of bitstreams: 1
description	provenance	Approved for entry into archive by Julene Vermeulen(julene.vermeulen@up.ac.za) on 2008-03- 14T06:03:00Z (GMT) No. of bitstreams: 1
description	provenance	Made available in DSpace on 2008-03-14T09:00:22Z (GMT). No. of bitstreams: 1 PHv1_Wilkinson-009.pdf: 145337 bytes, check

format	extent	145337 bytes	Generated by System
format	mimetype	application/pdf	



Checksums in DSpace



Checksum generated by Checksum software:

	•	After you save the results you will be able to check the integrity of your files. Press the <u>Save</u> (Ctrl+S) button to save.		
	Displa	ay the full path in file list		
Var	me 🔺		Size	Checksum\State
È	smithgr	oenewald_karm08.pdf	3,663 KB	E5A95F72B5F982A4B08B3C35DA7DC80C

Identical to Checksum generated by DSpace (UPSpace):

File	Size	File Format	Checksum	
smithgroenewald karm08.pdf	3,751,421 bytes	Adobe PDF (known)	e5a95f72b5f982a4b08b3c35da7dc80c	: (MD5)



AIP – Preservation Info – Fixity

Digital Signatures

- Digital signatures added to full text
- Compute a digital signature for digital masters & store signature in technical metadata of object
- Compute signature for complete item and store externally to repository

AIP – Preservation Info – Fixity







Persistent Identifiers

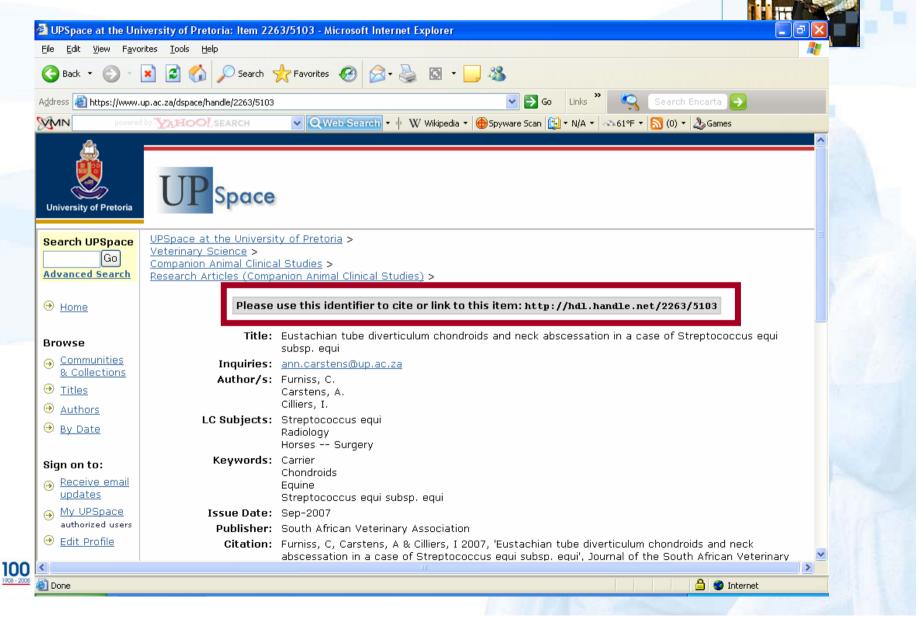
- Web references are untrustworthy; telephone numbers, IP addresses, Social Security numbers share properties of PID's – more trustworthy
- Persistent Identifiers: globally unique name assigned to a digital object that can be used in perpetuity, to refer to and to retrieve the digital object
- CNRI Handle System



AIP – Preservation Info – Reference



Persistent Identifiers



Deposit Licence



Coords C	c.za/dspace/submit	Google		
Google 💽 -	isity of Pretoria: UPSpace Distri	<u>à</u> • D	- 🖶 - 🔂 Page -	O Settings▼
	NON-EXCLUSIVE DISTRIBUTION LICENSE In order for UPSpace to reproduce, translate and distribute your submission worldwide your agreement to			~
	 the following terms is necessary. By submitting this license, you (the owner of the rights) grants to the University of Pretoria the non-exclusive right to reproduce, translate (as defined below), and/or distribute your submission (including the abstract) worldwide in print and electronic format and in any medium, including but not limited to audio or video. You agree that the University of Pretoria may, without changing the content, translate the submission to any medium or format for the purpose of preservation. You also agree that the University of Pretoria may key more than one copy of this submission for purposes of security, back-up and preservation. You represent that the submission is your original work, and that you have the right to grant the rights contained in this license. You also represent that your submission does not, to the best of your knowledge, infringe upon anyone's copyright. If the submission contains material for which you do not hold copyright, you represent that you have obtained the unrestricted permission of the copyright owner to grant the University of Pretoria the rights required by this license, and that such third-party owned material is clearly identified and acknowledged within the text or content of the submission. IF THE SUBMISSION IS BASED UPON WORK THAT HAS BEEN SPONSORED OR SUPPORTED BY AN AGENCY OR ORGANIZATION OTHER THAN THE UNIVERSITY OF PRETORIA, YOU REPRESENT THAT YOU HAVE FULFILLED ANY RIGHT OF REVIEW OR OTHER OBLIGATIONS REQUIRED BY SUCH CONTRACT OR AGREEMENT. The University of Pretoria will clearly identify your name(s) as the submission, and will not make any alteration, other than as allowed by this license, to your submission. All items in the UPSpace collection are subject to copyright. For more information on the SA Copyright, visit the SA Copyright Act No. 98 of 1978 (as amended) available at http://www.buys.co.za/publications/cyberlaw/CopyrightAct.htm 			
	I Grant the License			
1		Second Second		100% -

Deposit Licence



Permissions given to the repository

The IR shall distribute electronic copies of the work for the lifetime of the repository, or based upon an agreed time span, and translate it as necessary to ensure it can be read by computer systems in the future.

The Repository's Rights and Responsibilities

The IR may electronically store, translate, copy, or re-arrange the work to ensure its future preservation and accessibility, unless notified by the depositor that specific restrictions apply.

Source: SHERPA Project Document – Report on a deposit licence for E-prints



Storage Management

- Storage hardware is a key component of a repository
- SAN (Storage Area Network) vs NAS (Network Attached Storage)
 - Increased scalability: up to 16 million devices can be added
 - All other participants on SAN can connect and see each other
 - High-speed throughput: carry traffic between devices at 2 Gb/s
 - Independent of other network operations functions separate from any LAN



Preservation Policies & Tools



- erpaTool Digital Preservation Policy Tool
- Cornell Digital Preservation Tutorial
 <u>www.icpsr.umich.edu/dpm/dpm-eng/contents.html</u>
- DRAMBORA <u>http://www.repositoryaudit.eu</u> Digital Repository Audit Method Based on Risk Assessment Toolkit





Institutional Repository Workshop

A to Z of digital preservation within an Institutional Repository Business Plans, Policies, Digitization, Metadata, Implementation, Marketing & Buy-in and many more ...

> **1 – 3 October 2008** University of Pretoria

www.library.up.ac.za/irtoolbox/workshop.htm

OR

E-mail:

ria.groenewald@up.ac.za ina.smith@up.ac.za





Join our IRSpace CoP!

E-mail us:

ina.smith@up.ac.za



Will your work withstand the test of times to come?

Questions?

Sources

- 1. Caplan, P., and Guenther, R. (Summer 2005). Practical Preservation: The PREMIS Experience. Library Trends, Vol. 54, No.1
- 2. Data Migration (Wikipedia). <u>http://en.wikipedia.org/wiki/Data_migration</u>. Retrieved May 2008
- 3. DIG35. http://xml.coverpages.org/dig35.html
- 4. Digital Preservation Research. http://www.kb.nl/hrd/dd/dd-en.html
- 5. Dioscuri, KB Nationaal Argief. http://dioscuri.sourceforge.net/ Retrived May 2008
- 6. Draft Charter on the Preservation of the Digital Heritage, Item 8.6 of the provisional agenda. 19 August 2003. General Conference 32nd session, Paris
- 7. Emulation http://www.kb.nl/hrd/dd/dd_projecten/projecten_emulatie-en.html
- 8. Metadata Framework Working Group. <u>http://www.oclc.org/research/projects/pmwg/pm_framework.pdf</u>. Retrieved April 2008
- 9. Migration <u>http://www.kb.nl/hrd/dd/dd_projecten/projecten_migratie-en.html</u>
- 10. Mullins, J.L., Allen F.R., and Hufford, J.R. (April 2007). Top ten assumptions for the future of academic libraries and librarians: A report from the ACRL research committee. C&RL News, Vol. 68, No. 4
- 11. NDHA Business Requirements Specification, July 2005 www.natlib.govt.nz/ndha
- 12. NISO. <u>http://www.niso.org/kfile_download?pt=RkGKiXzW643YeUaYUqZ1BFwDhIG4-</u> <u>24RJbcZBWg8uE4vWdpZsJDs4RjLz0t90_d5_ymGsj_IKVa86hjP37r_hM9t9qad1BrrORLqssvegis%3D</u>. Retrieved May 2008
- 13. OAIS Reference Model for an Open Archival Information System http://public.ccsds.org/publications/archive/650x0b1.pdf
- 14. PLANETS http://www.planets-project.eu/
- 15. PREMIS Working Group. http://www.oclc.org/research/projects/pmwg/presmeta_wp.pdf. Retrieved April 2008
- 16. PRONOM http://www.nationalarchives.gov.uk/pronom/ Retrieved May 2008
- 17. Rog, J., van Wijk, C. (2008) Evaluating File Formats for Long-term Preservation, National Library of the Netherlands; The Hague, The Netherlands. http://www.kb.nl/hrd/dd/dd links en publicaties/publicaties/KB file format evaluation method 27022008.pdf
- 18. Weinberger, E., Clayton, R., Anderson, R. A security policy for a digital repository. Cambridge University. http://www.cus.cam.ac.uk/~ew206/security-policy.html
- 19. Wikipedia on Z39.87. http://en.wikipedia.org/wiki/Z39.87
- 20. Wikipedia on NISO. http://en.wikipedia.org/wiki/National_Information_Standards_Organization
- 21. http://www.nyu.edu/its/humanities/ninchguide/VI/ Retrieved March 2007

