STELLENBOSCH UNIVERSITY LIBRARY SERVICE SYMPOSIUM, 27-28 OCTOBER 2005

TECHNOLOGY PEOPLE COME FIRST: HOW CLIENT-CENTRED ARE OUR LIBRARIES?

eResearch - Opportunities for Researchers and Challenges for Librarians, or vice versa? Roy Page-Shipp.

# Theme

Global connectivity via the WWW is changing the ways in which researchers not only communicate their findings to each other, but collect and share data and information in a variety of media, and share computational capacity, models and other tools Libraries have supported researchers' earlier needs, but they have new needs What? So What? .....Now What?

Based in part on SARIS - Ford Foundation Project

**Global Connectivity - Greater speed** and Scope of communications  $\succ$  The War in your home, daily: Baghdad or Ellis Park! > Ma-in-law (or granny) on e-mail – daily! Online banking, news, bookings, shopping > e-everything – learning, research, content > Wi-fi, VOIP and other formats > - and you can 'get it all on the Web' -Google moves into indexing content

# ..and Cellphones



# And Cellphones

PCs 130M pa, Cellphones 600M pa and rising

- Cellphones in SA 19 M users in June 2004, access for 71% of the popn, 90% new connections prepaid, SMS traffic up 1000% in 2 years
- Teenager to teenager even <u>at school</u>
- 'Find me' facility,
- Direct Medicines, FNB, Blood bank, Air Rewards
- > Catching murderers
- Photos and videos
- The Blackberry



## Transfer Rates: Bytes & Gigabytes





#### PRAE ANY 72, 23





a la la la la la la



2005 9 24

## Transfer Rates: Bytes & Gigabytes

2pp e-mail	30 000 bytes	1/300 KJV	
1200 words	30 kilob		
Digital Photo	150 -5 000kb	1/66 – 1/2 KJV	
King James	10 000 000 b	1 KJV	
Version Bible	10 Mb		
	Max Rate	Time per KJV	
Telephone Line	64 kb/s	2.5 min	
ADSL Line	512 kb/s	20 sec	
I Burst	1 Mb/s	10 sec	

#### What makes us human? Experimental data Completed Work – Scholarly Discourse



# **Global Scholarly Discourse**

- Researchers communicating with their peers
- Print publication became the vehicle as the world 'grew'
- Publishers threatened to take control of the discourse – Copyright, costs etc
- Now the WWW enables (almost) real time discussion with counterparts (almost) anywhere
   So publish via the WWW a no brainer!
   ...or is it?

### What do researchers (clients) need?

- Registration, ...claims of precedence for a scholarly finding.
- Certification, ,, establishes the validity of a registered scholarly claim.
- Awareness/visibility, ...allows actors ... to remain aware of new claims & findings & ... eligible for int'n'l collaboration in the 21st C
- Archiving, this preserves the scholarly record over time.
- Reward, ... for .. performance ..based on metrics derived from the system. based on Roosendal & Geurts
- and Ownership of the IP?

# Open Access – the facts?

- Costs
- Speed
- > Visibility/Harvesting
- Curation/preservation
- Citation
- > Peer review
- Every one Complex
   See The Facts about Open Access, Kaufmann-Wills Group LLC, 2005

	Journals		Open Access				
	Print	Online	Gold	Green			
Costs	Subs	Subs	Author, Sponsor	Institut'n			
Speed of Registration	On Subm	On Subm	Immediate or delayed	Immediate			
Visib'ty/Harvest'g	Good	Good	Variable	Unreliable			
Curat'n/Presrvt'n	Good	Good?	Dodgy	Dodgier			
Citation	Good	Good	Better	?			
Peer Review	Conventional		Some Innovation	Institution'l			
Reward	DoE List		Possible?	Unlikely			
Fin Sustainability	Budget Threat		Dodgy but improving				
Some partially valid generalisations in a very complex area							

## **Open Access & Librarians**

"The Library role in Open Access is Leadership,

Service,

Standardisation"

John Shipp - University of Sydney

# ..and Google Scholar!

"I can find it myself" Last Resort – order it from the **British Library!** >About R120/copy, not Transferable!! Protect the Patch or Lead and Train?

# myGrid - UK

- > Imminent 'deluge' of data
- > Highly complex and heterogeneous data
- Convergence of data and literature archives



# **Digital Curation**

"The active care and management of datasets for their scholarly useful life" No more 'Dark Ages' > Data collection is expensive Marketing and Promotion > New career opportunity for librarians > Metadata attribution  $\equiv$  Cataloguing Leadership, Service, Standardisation?

#### Librarians and Digital Curation

Metadata attribution 
 Cataloguing/Indexing

 Do it and train others to do it
 Learn to market datasets 
 SDI

 Take over the longer-term responsibility
 Make yourselves indispensable, counter threats of redundancy



In SA - the SANReN



**OCLC view of New Information Flow in Scholarly Communication** 



'eScience is about global collaboration in key areas of science, and the next generation of infrastructure that will enable it.' (includes social science data sets): John Taylor - DG Research Councils, Office of S &T, UK

Purpose is to allow scientists to do faster, different, better research

<u>Prof Tony Hey</u>: Core programme director for UK eScience and chair of JISC committee (£220M over 5 years)

# SA Scene – eScience

- Some researchers (eg Human Genome) battling with global data transfer – VM reporting real threat of isolation
- Use of data and models across research teams is increasing
- SANReN is here; linking SA to GéANT, Internet 2, AREN; still too costly
- Some potential users will cope, others need help – need for a national Helpdesk
- Science a new growth area SKA



# DAME Technologies & Collaboration

AURA - Advanced Uncertain Reasoning Architecture – high performance pattern matcher developed by University of York and Cybula Ltd.

QUOTE – "On The Engine" intelligent engine signature collection / local diagnosis system developed by University of Oxford for Rolls Royce / DS&S.

Decision Support – University of Sheffield.

GRID architecture / web services – University of Leeds.

![](_page_24_Picture_0.jpeg)

SA Government focus on Astronomy SA well placed - Clear Skies – wide open spaces Southern Hemisphere – more interesting view of the Universe – Hence SALT Managed Radio Interference

![](_page_25_Figure_0.jpeg)

Why Radio Astronomy?

- Radio Waves because of Cosmological expansion
- Frequencies 30 MHZ and 40 GHz, ie wavelengths from 10 m to 7 mm.
- > Waves penetrate interstellar & terrestrial dust

![](_page_26_Picture_0.jpeg)

**Objectives** 

Events before the formation of the stars ~ 13,5 Bn years ago

- Very faint signals, lots of 'noise' hence a big antenna – 1 sq Km
- Flat array on the ground too big for a dish

Computational beam-forming

![](_page_27_Figure_0.jpeg)

![](_page_27_Picture_1.jpeg)

Cost 10-12Bn US\$, mostly from overseas
1000's of antennae
Each producing ~ 3Gb/s = 300KJV/s
Overall 300 000 KJV per second!
Start 2012?
KAT Pilot Project – 2004 - 2006

# Components of eResearch

eScience		Digital Curation &	Access to eContent					
Data Transfer	Tools & applications	Primary Data Sharing	Preservation	Commercial Publishing	Open Access Publishing			
by definition this is								
Science employing transfer and sharing of large volumes of data	Software that allows manipulation, modeling and analysis of data	Making research data available to other researchers	Active management of databases including promotion of effective and widespread use of the datasets for their scientific & scholarly useful life	Contribution to & use of published resources requiring payment by readers	Contribution to & use of published resources where content is regarded as 'free'			
which requires								
Access to remotely held large datasets & high performance computing via affordable high bandwidth	Access to models, source code and open standards	Accessible repositories & quick reference	Preservation & curation repositories & access mechanisms, archival skills & infrastructure	Affordable licenses for researcher access & discovery mechanisms	Serviceable infrastructure for publication and access			

Researcher Requires: Perpetual access, Curation, Training, Marketing

Supplier must ensure: Security - Access, Authorization, Authentication

#### A 'Team SA' approach for eR3SA

**eResearch Board** 

Governance & Management Model

eResearch Development & Innovation

Function

eResearch Service Delivery

#### **Future eResearch activities**

•Web Access Framework – eResearch Portal

•Data Transfer and Sharing (processes and protocols, 3As, helpdesk)

•Open Access (Standards, common software, institutional repositories)

•The eResearch Librarian (Training and re-orientation)

•Digital Curation Services (Standards, software, marketing & training services)

**Lead Users Forum** 

![](_page_29_Figure_13.jpeg)

Usually sub-contracted to competent agents in the system

# Now What? – for Librarians

- ➢ New Challenges ≡ New Opportunities
- Precious Basic Skills
  - Organisation and protection of valuable information
  - Ontology creation
  - Use of complex search tools
  - Metadata attribution
  - Marketing and training
  - Service orientation
- Common IT base
- Librarians or Someone Else?