Positioning ETDs in the eResearch arena: a South African case study

Monica Hammes

University of Pretoria : Academic Information Service

monica.hammes@up.ac.za

8th International Symposium on Electronic Theses & Dissertations

Sydney, 29 September 2005

Outline of Talk

- What is eResearch and why does it matter?

- Challenging the current scholarly communication system

- eResearch, scholarly communication and etds

What is eResearch and why does it matter?

What is eResearch?

eResearch   **eScience**   **cyberScience**

Scientific endeavours that are enhanced by **ICTs** and an **abundance of data**

Leading to new research practices that are highly collaborative, network-based, data-intensive and conducive to the creation of knowledge environments with the capacity for unparalleled global collaboration

- Use and re-use of data

- Knowledge communities

It is dependent on....

- Computing capacity and capability (grids, supercomputers, clusters, workstations)

- Mass storage

- Data capturing and data transmission infrastructure to share large data-streams, datasets and models (NReNs)

- Powerful data mining techniques

- Networking (including optical, wireless, ubiquitous, ambient)

It is dependent on….

- Digital libraries/data bases with sufficient metadata for potential users to find the data and be satisfied of its value and provenance

- Software (operating systems, middleware, domain specific tools/platforms for building applications, analysis, visualization)

- Services (education, training, consulting, user assistance)

- Access to the global research literature and infrastructure for open access publication

Examples

- Bioinformatics

- Astronomy

- Earth observation

- Oceanographic studies

- Environmental pollution monitoring

- Advanced engineering
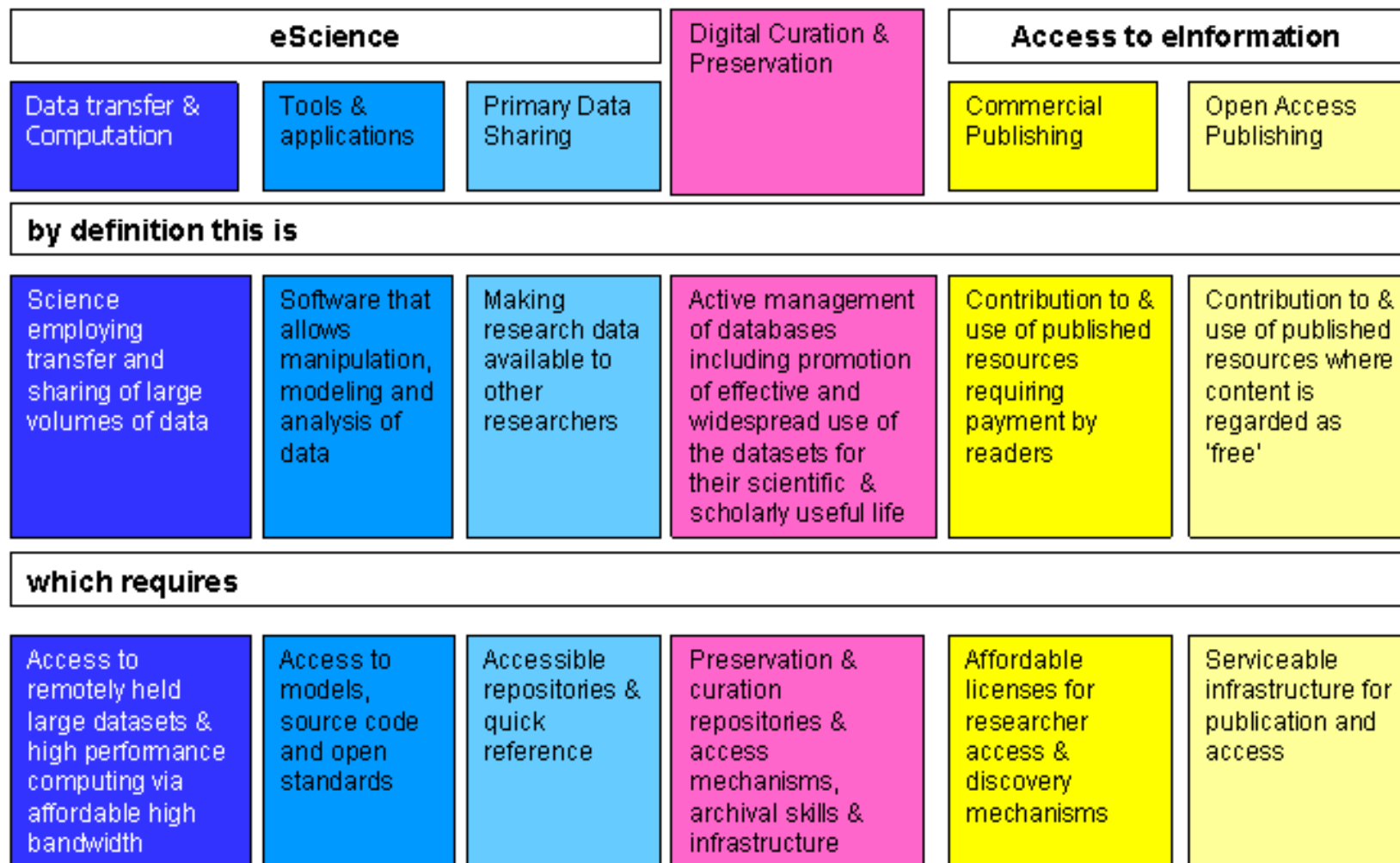
- Population research

The South African SARIS Project

Background

Researchers in developing countries are faced with dramatically improved opportunities for global collaboration but also stand the risk to be left out of the modern research milieu due to inadequate infrastructure

- Investigation into the declining affordability of access to global research literature

- eResearch presents a broader range of support challenges in need of a coherent solution
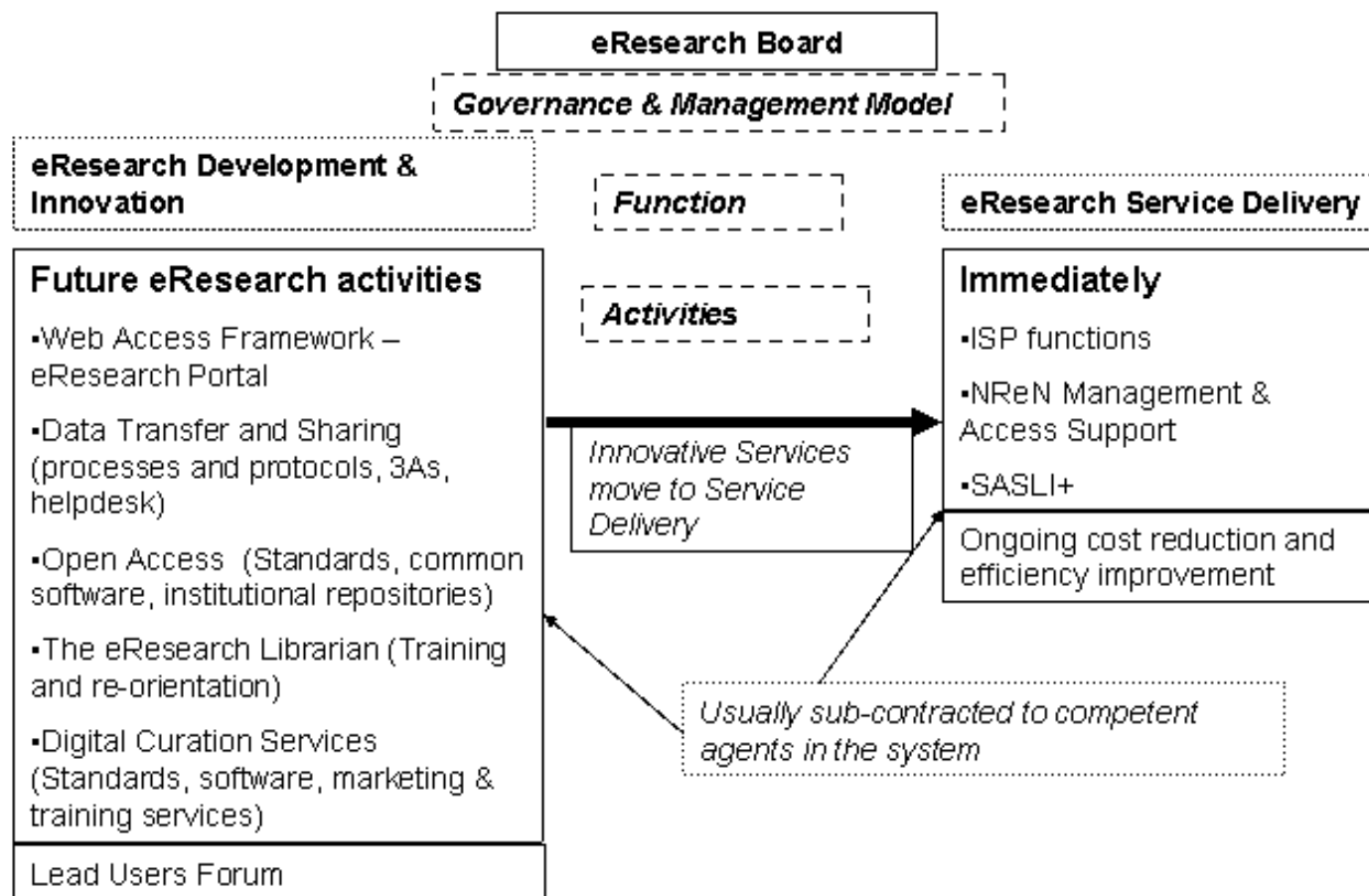
Components of eResearch

eScience			Digital Curation & Preservation	Access to eInformation	
Data transfer & Computation	Tools & applications	Primary Data Sharing		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

eRS3A: Team SA approach: pre-competitive

SARIS Project Outcome: Proposed Structure for eResearch Service for SA

eResearch Board

Governance & Management Model

eResearch Development & Innovation

Function

eResearch Service Delivery

Future eResearch activities

Activities

Immediately

- 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

Innovative Services move to Service Delivery

- ISP functions
- NReN Management & Access Support
- SASLI+

Ongoing cost reduction and efficiency improvement

Usually sub-contracted to competent agents in the system

Challenging the current scholarly communication system

Dramatic changes in the nature of scholarly research require corresponding fundamental changes in scholarly communication.

Terhorst (2005)

Journals are obsolete formats; they must evolve to accommodate data-intensive science

Djorgovski (2004)

The learned article should be instead regarded as more of a functional tool, to be used with the appropriate combination of software based processing and transformation of its content

Rzepa & Murray-Rust (2003)

The peer-review system does not take full advantage of new possibilities

Cronjé (2005)

New units of communication, including datasets, simulations, software as well as complex documents consisting of multiple data streams should be accommodated
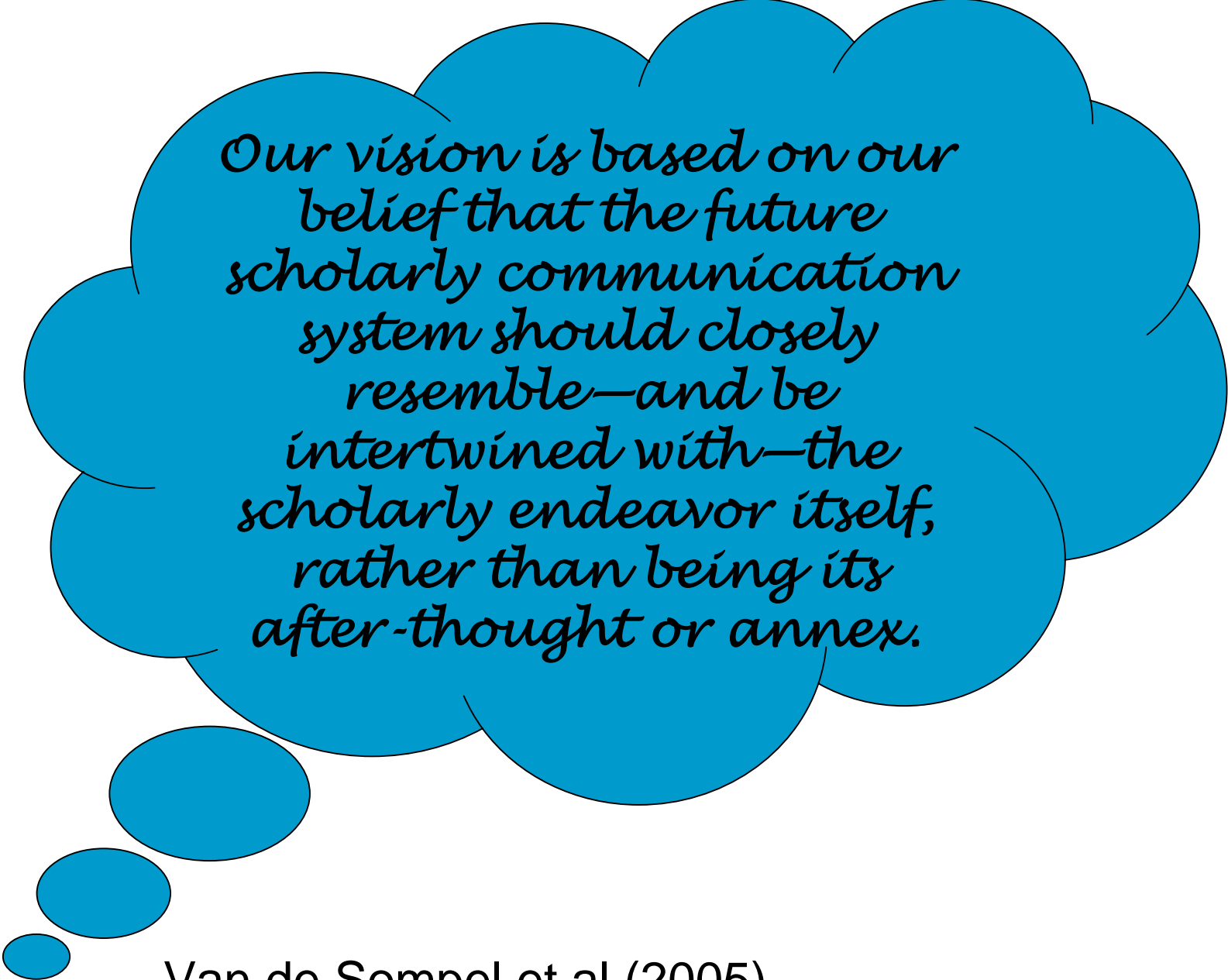
Van de Sompel et al (2004)
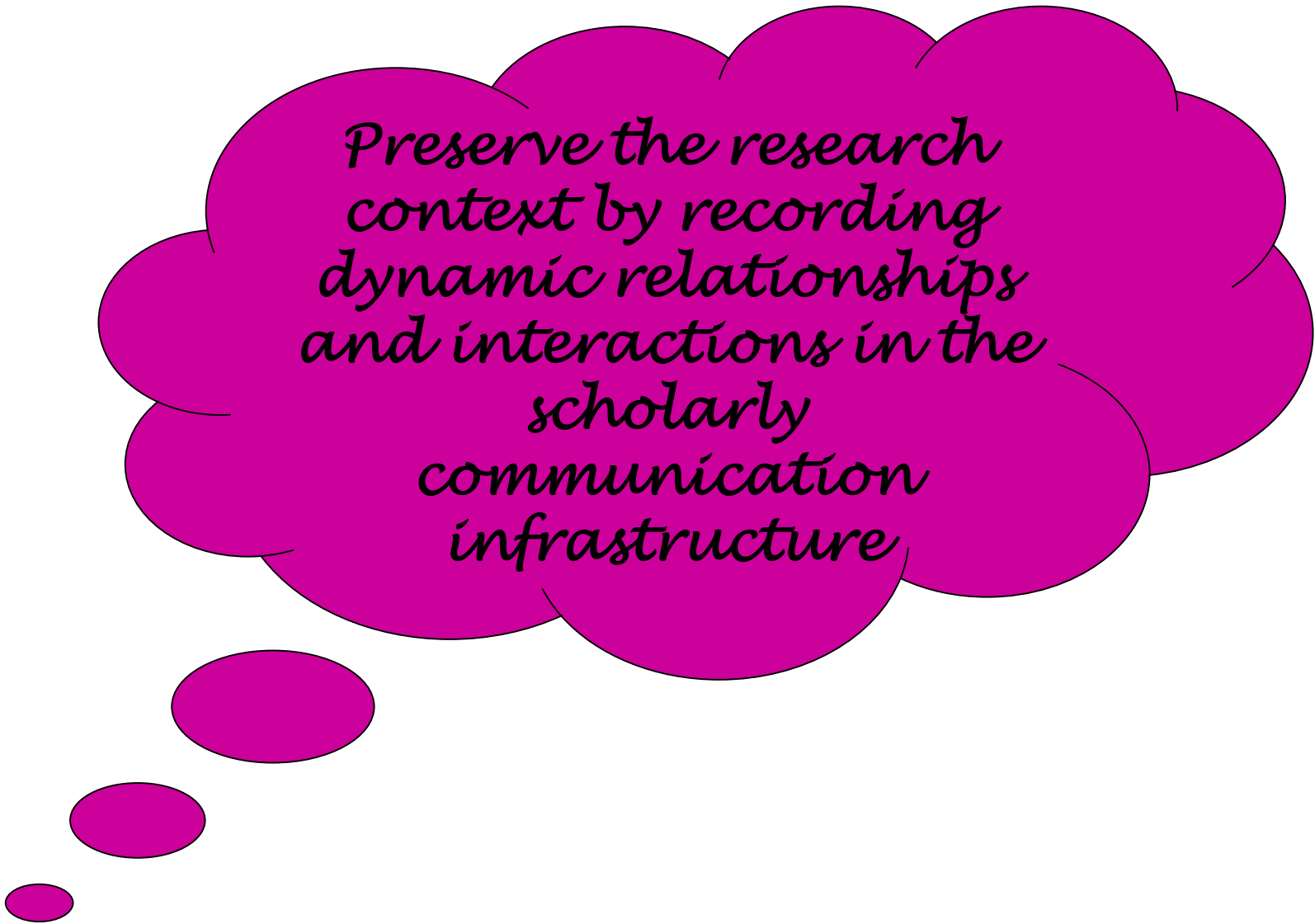
Preserve the research context by recording dynamic relationships and interactions in the scholarly communication infrastructure

Jeffery (2005)

Develop new metrics to assess the quality of scholarly assets and for the evaluation of the performance of actors in the scholarly system
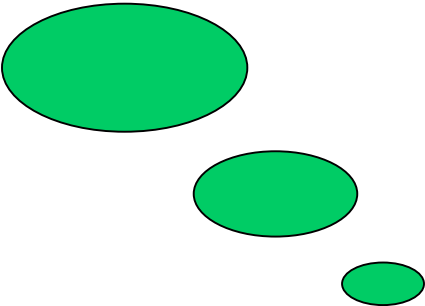
Rzepa & Murray-Rust (2003)

Our vision is based on our belief that the future scholarly communication system should closely resemble—and be intertwined with—the scholarly endeavor itself, rather than being its after-thought or annex.

Van de Sompel et al (2005)

eResearch, Scholarly Communication and etds

..it may not be simply the technical ability to reproduce and distribute articles electronically (e-publishing), but also the emergence of highly collaborative, large scale investigations and analyses (e-science) that is likely to lead in the field of scientific communication and significant changes in the way such communications are produced, curated and dissemminated

Lucier, 1990

Research Portal

Interface for service delivery in a personalized manner:

- Components integrated in an organisation's portal, or

- accessed from any Internet service point - isolated researchers, or those in poorly resourced institutions

Research Portal Attributes

[1] A single access point to a family of repositories for data, digital objects and publications

[2] Online information resources available by affiliation: commercial and open access with alerting services and federated searching capabilities and a pay-per-view facility for any other resources

[3] Facilities for Communities of Practice/Curiosity

[4] Assistance in submitting large databases/streams to SANRN

[5] Online Research Support Tools

[6] Single sign-on, authentication, authorization

Portal enhancements for post-graduate research

An integrated environment that will deal with everything between the registration of a post-graduate project up to an etd on the web. Elements to be included would be

[1] Shared work space for students and supervisors

[2] All university requirements available at the point of need

[3] Guidelines and tools for postgraduate research and thesis writing

[4] Links and functionality related tot funding

Portal enhancements for post-graduate research

[5] A work flow that will keep the project on track from registration to a final etd on the web

[6] Information on ethics and a plagiarism detection service

[7] Referencing and bibliography building software

[8] Discipline specific tools: polling, surveys, data processing

[9] A management process that will keep track of publication activity per department/university

4 Issues for further exploration

- Data curation

- Intellectual property: data, simulations, programmes

- Preserving the context of research

- Examination (peer review)

Conclusion

• Graduate research is training to be fully-fledged researchers

• Universities are well positioned to try out new ideas and could start these changes

• The ETD community is well placed to create awareness for these issues

Thank you!

monica.hammes@up.ac.za