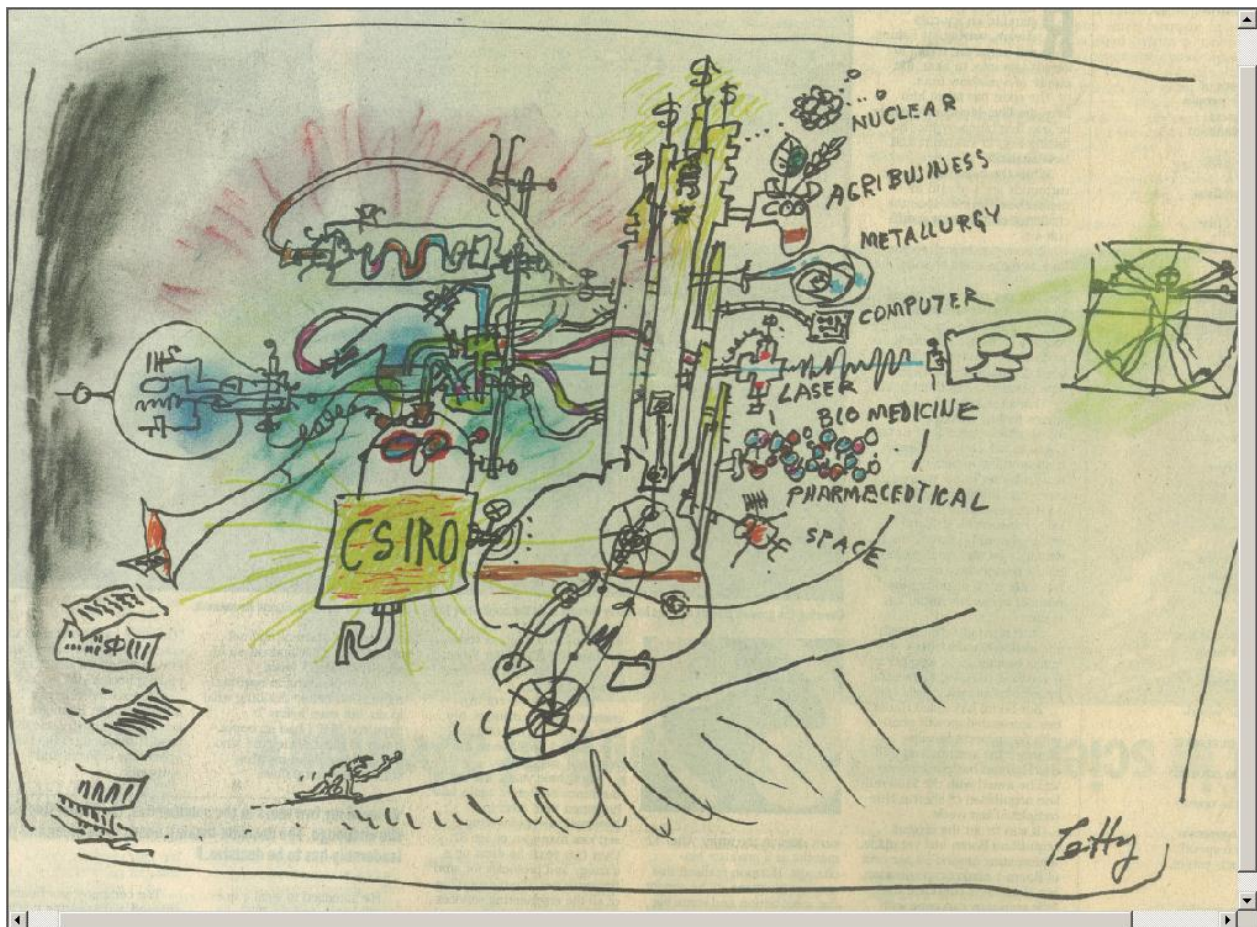


Information Futures Commission

eScholarship Research Centre Submission – Number 1  
9 May 2008

Gavan McCarthy  
Director, eScholarship Research Centre



‘Everyday, my PC background (courtesy of Bruce Petty and *The Age*) reminds me that the drivers and vested interests of contemporary Australia, while lofty in their vision, risk leaving behind only a meagre trail of evidence but more importantly the people, the systems are meant to serve, are left struggling to remain connected.’

Gavan McCarthy, Director, The University of Melbourne eScholarship Research Centre

## **Introduction**

This submission, prepared by Gavan McCarthy, Director of the eScholarship Research Centre, references a number of documents prepared by Centre staff. We have lodged these related submissions separately. This submission acts as an overview document and reflects discussions and debates within the Centre during the course of the Information Futures Commission. That said, this document is the work of the Director and reflects his particular views and ideas.

Joanne Evans has prepared a substantive scholarly piece that looks at the background to scholarly practice in the digital age and captures a range of key considerations that should influence the thinking of the Commission. Simon Porter has written a focussed paper on how we can better use information about research and researchers to both aid research practice itself as well as the management and reporting of University research activities. Eve Young has submitted a paper that looks at the role of ePublishing (what we do with published outputs) and how digital and networked technologies influence this. Lyle Winton has dealt with what he has called the ‘Information Journey’ and explores various issues including the emergence of service oriented architectures.

We established the eScholarship Research Centre (ESRC) in 2007 specifically to address, in a very practical way, the challenges of scholarly practice in the digital age – the underlying theme of the Information Futures Commission. The ESRC actively collaborates with researchers to help them understand the challenges posed by new technologies and to find sustainable scholarly practices that will ensure that their contribution to the corpus of human knowledge is recognised and utilised over time. The ESRC also works within the University to help build an understanding of the information infrastructure necessary for the University to function effectively and efficiently in a largely digital scholarly world. By information infrastructure we mean both hardware and networks (ICT infrastructure), and the role of persistent well-structured information objects (Persistent Related Information infrastructure or PRI infrastructure).

The ESRC draws on expertise base that includes archives, records, historiography, physics, English, design, art history, computer science, and history and philosophy of science. However, this diverse group is united by world-leading expertise and experience in contextual information management and an operational model that is based on direct collaboration (as research colleagues) with researchers and research projects within the University and with external bodies both in Australia and overseas. Our own research agenda could be simply characterised in the question: How can well-structured contextual information frameworks in open network environment be utilised to make ‘stuff’ meaningful?

## **Issues, Assumptions and Quick Comments**

The breadth of the Information Futures Commission and the issues it raises for scholarly information in the digital age are complex and wide ranging. This submission will in no way attempt to cover all possible issues in depth. Below is a list of assumptions, brief comments and issues that while not expanded or articulated in detail provide an overview these broad issues from our perspective.

## **ICT Infrastructure**

- Data Storage – it is taken as given that University must provide a centralised data storage service. Demand for this has been growing steadily over the last decade but is now facing an exponential growth in demand. We are aware that processes to establish this service are advancing, however, this process needs to be accelerated and appropriate operational and associated funding models developed that suit the funding and processes of research.
- Virtual Machine (servers) linked to data storage are the front-end servers and services that enable data to be utilised and the same arguments for storage apply to VM.
- Network connectivity and bandwidth – this is the third element of the ICT triumvirate that we must address immediately for researchers and research projects to function effectively and efficiently. This third element could also include wireless and digital telephony services. While development and installation projects are underway, they are coming late on the scene for a ‘leading’ University and we must expedite them as a matter of urgency.

NOTE: Policy must be developed and implemented to enable those with demonstrated needs and capability to lead in the development and deployment of these services.

## **Library Infrastructure**

- The ‘book – digital’ debate. In the digital – network world, the primary form of just about all information objects will be the digital form. However, the ability to create, as required, analogue or print forms will become increasingly localised. Print versions of information will maintain their role as a human-preferred manifestation in many circumstances but the University has to recognise that this form is not the primary publication form, or indeed the primary management form of the object.
- Libraries will continue to have a critical place in the University’s built and scholarly service infrastructure and there must be investment in buildings, storage, reference and usage services.
- Library cataloguing and reference services, some of which are very efficient, reflect the book paradigm and will/are only partially successful in the digital world. The University must acknowledge this and put processes in place to move beyond the current state. This will require research innovation and testing.

## **High Performance Computing Infrastructure**

- High profile discipline-based research projects and communities have the ability to attract significant funds to build what become essentially siloed services and capabilities that meet the needs of particular communities of practice. However, experience has shown that generally this does not lead to the extension of those services and capabilities to the long-tail of University researchers (that is most of them). The needs of the long-tail are becoming critical and many researchers risk becoming dysfunctional, disconnected and unproductive unless there is a broader sharing of this expertise and capability.
- High performance computing for analytical, visualisation and other computer intensive research activities is becoming widespread across all disciplines, and this includes the much ignored arts, humanities and social sciences.

## Teaching and Learning

- If the University of Melbourne is going to be a research-led teaching and learning University then all of the above issues apply. This submission will not address the specific issues facing teaching and learning in this submission. However, if the university community of scholars (students and staff) are going to be functional and efficient in the digital scholarly environment then there needs to be concerted effort to develop and deploy instructional courses at all levels – from first year entrants to re-skilling programs for senior academics.

## People

- People are both the cornerstone and the heart of the scholarly information industry – after all, it is all about human knowledge. We need to ensure that all staff and students have the opportunity to engage in what will be the biggest skill overhaul in the history of scholarly practice. The introduction of the printing press, regarded historically as revolutionary, will seem staid and cautious when compared to the rapid and disruptive introduction of networked digital technologies.
- In a fiscal environment that is highly constrained for all University staff and for many students, there is precious little capacity for re-skilling. While it is difficult to see how the University will address this in the short-term, perhaps the simplest solution is that all position descriptions include digital scholarly re-skilling and training (in both informal and formal modes) as a core element.

## Cultural Collections

- One of the University's greatest assets is its cultural collections that cover a vast array of materials – from archives to artworks. However, the University also has born-digital cultural collections, many of which are the results of research projects, for example, the Australian Women's Archives Project *Women Australia* register; and the high profile *Bright Sparcs & Australian Science at Work*. For a fuller list of these types of born-digital cultural assets that are connected to the ESRC see: [www.esrc.unimelb.edu.au/digipubs/resources.html](http://www.esrc.unimelb.edu.au/digipubs/resources.html).
- We need to change our thinking at the highest levels to acknowledge these types of cultural assets and to locate philanthropic funds to support these public good knowledge transfer resources.
- The University of Melbourne Archives is a nationally recognised repository of traditional 'paper-based' social and community archival materials. The location of the main repository off-site appears to have led to isolation of the professional archival staff from general University life and in particular from the intellectual and conceptual strivings of similar professionals and academics in other parts of the University. We should address this issue with some urgency.
- The shift to a scholarly environment that is fundamentally digital is bringing to the archival world a new cohort of researchers and research support staff traditionally removed from the archival process. The profoundly complex issues emerging in the eResearch data management space relationship are essentially archival in the fully professional and idealistic sense of 'archive' not the common misconception of what archives are in the IT sense (just a copy) or in the caricatured sense of dusty and

forgotten. It appears that the best place to address these issues is through the eResearch data management space but traditional archivists and cultural collections managers must have the ability and desire to engage in this space and acknowledge that their current practices are not sufficient for the archiving of digitally-born or indeed digitised traditional materials.

## **The eScholarship Research Centre and the post Information Futures Commission world**

The Information Futures Commission in its draft contemplations identified:

- An aspiration – We will innovate, collaborate and influence to position Melbourne as a leader in the application of scholarly information and technologies to advance scholarship, investing in areas of research strength.
- A 2012 goal – Melbourne will be a leader and innovator in information science. Library and information professionals with deep knowledge will be available to work with researchers in all areas of research focus.
- A 2020 goal – We are recognised internationally as a leader in the application of scholarly information and technologies to advance research and scholarship.

Sitting under all these aspirations and goals is a presumption that the University will be actively engaged in research into scholarly practice in the digital age and will contribute to the world of digital scholarly research through innovation, testing, collaboration and implementation in a global environment. The eScholarship Research Centre is a proto-centre of these activities but is currently critically under-funded to fulfil this mission on the scale required for the University. The Information Futures Commission provides an opportunity to define a new operational, fiscal and governance space in which the ESRC in combination with the closely related work of the Director of eResearch and other eResearch and research data management initiatives can help the University achieve these objectives.

### **Current Constraints**

Most in the University would be unaware that the relative small operating budget for the ESRC is less than one third funded from central University sources. In order to keep some of the best and brightest informatics practitioners, thinkers and researchers connected with the University, we ‘sell’ their services to other universities and organisations. The opportunity cost to the University of Melbourne is significant and is constraining internal development and advancement. There is a positive side to external collaborative engagement but it should not be to the detriment of internal development and services to our own researchers, archives, cultural collections and teaching and learning.

The future of the eScholarship Research Centre and the opportunity it offers the University of Melbourne should be a primary consideration of Commission.

### **Reflections of a more Scholarly Nature**

Dr Joanne Evans’ submission, which she is refining for full academic publication, eloquently captures some of the fundamental issues. What follows is series of selected quotes from her submission to highlight critical issues:

### **On Joanne herself . . .**

‘I am both a scholar and an archival system developer, with both roles closely entwined. My information systems development work, which is guided and supported by a strong theoretical framework, in turn illuminates and develops that framework. With much to learn about the informatics of the digital and networked age, I believe that such an approach fosters innovation in both research and practice. I consider that this is a must as we move from just applying ICT to automate existing processes and systems, to re-conceptualising, re-figuring, re-engineering and transforming them, in some cases beyond recognition, in order to fully exploit the capabilities of new technologies. There is much to discover, learn and understand about the new technologies and their application, as we transition from a paper to a digital world.’ (Evans 2008 Introduction paragraph 4)

### **On Web 2.0 . . .**

‘Web 2.0 is an architecture of interoperability, participation and collaboration that is reflected in the way applications and infrastructure are constructed, as well as the way they are utilised. The attraction lies in the potential for productivity improvements and in the new capabilities that can be realised, with software conceived as a service rather than a product, and is able to be embedded in information interfaces.’ (Evans 2008, The Next Ten years paragraph 3)

### **The National Library of Australia as an exemplar . . .**

‘In March 2007, the NLA published their *IT Architecture Project Report* which lays out their aim of embracing wholeheartedly a service oriented, single-business approach to the provision of digital library services, including the adoption of open-source solutions where ‘these are functional and robust’ (National Library of Australia 2007). There are many interesting and exciting aspects of the NLA report, and I would suggest it is a ‘must read’.

- One of the outcomes I would like to see from the IFC is the development of a similar plan for the University.
- I would also like to see the University embrace a service oriented approach to its information and administrative systems, and begin to create an environment which supports this approach.

I would like to highlight two aspects of the NLA’s report which I feel are important to achieving their goals. Firstly there is an emphasis on the need to change mindsets around how the library goes about acquiring and developing the technology it requires to support service provision, rather than on undertaking organisational re-structure.

A physical restructure is probably needed less than a new way of sharing ideas, communicating what is happening across the Library and building up the IT literacy of all staff. The single business approach provides a way of doing this by bringing people together to work on solutions to shared problems and by enabling all staff to be involved in testing and evaluating prototypes. (National Library of Australia 2007, p. 13)

Secondly peer review, iteration and experiential learning is noted as an important part of the strategy.

There will need to be more planning and peer review at the start of projects to determine what services are needed and how they might need to be enhanced to support the new requirement. Services will be built iteratively, with early versions only delivering the functionality that is immediately required. Short development timeframes for work packages will allow for prototyping, frequent business and user testing and experiential learning. (National Library of Australia 2007, p. 13)

(Evans 2008, Web 2.0 and Digital Library Services, paragraphs 1-3)

### **Furthermore . . .**

‘It goes to the heart of the NLA maintaining its visibility, viability and relevance in the digital and networked world of the 21st century.

- To be a leader in this area, the University should aim to participate in this community of practice in a variety of ways. Individual, project and institutional connections to this community will need to be made, encouraged and co-ordinated.
- The University should also try to foster within its own walls the same kind of commitment and interest in ‘interoperable, standards-based solutions’ to its information services, and draw together its own community of research and practice.’

(Evans 2008, Web 2.0 and Digital Library Services, paragraph 5)

### **Going forward . . .**

‘In practical terms, the University has to encourage *best practice* use of existing systems, at the same time as fostering the development of *better* systems for the future. Being able to identify and acknowledge both the capabilities and limitations of extant systems and new technologies will be important, as will seeking to involve all direct and indirect stakeholders in system re-development activities. The University needs to honestly appraise its activities, be able to learn from past successes and failures, and sponsor research and innovation in this area. It is a tricky balancing act between ensuring that we make the best of current frameworks, operational systems and tools, while at the same time seeking to realise their next generation.’ (Evans 2008, Other records of scholarly activities, paragraph 5)

### **Scalable Scholarly Information Infrastructure**

Borgman identifies the scalability of information management practices as one of the problems the data and information deluge of digital and networking technologies brings to the fore (Borgman 2007, p. 6-7). (Evans 2008, Scalable Scholarly Information Infrastructure, paragraph 1)

### **Sustainable Scholarly Information Infrastructure**

Borgman argues that it is the information not the IT that is ‘more crucial to scholarship’ and that we must now look to build a scholarly information infrastructure that ‘can outlive any technical architecture’ (Borgman 2007, p. 227). (Evans 2008, Sustainable Scholarly Information Infrastructure, paragraph 1)

### **Harnessing digital and networking capabilities**

‘What must be kept to the fore is that it is a digital *and networked* age that we are moving into. Thinking digital will only get you so far. Thinking network is what can turn the digital stuff into really, really useful stuff. In the first instance network infrastructure allows digital information to

be accessed from the desktop. Basic information infrastructure that comes from automating the access tools of the paper world combined with automated processing techniques can make it discoverable. But if we start to think about the digital stuff as networked information objects and recognise and invest in them as information infrastructure then the possibilities and potential become really interesting.’

‘I see that the role of information management professionals should be directed towards scholarly processes rather just at the end products. I feel that they have a role to play in designing and developing information management services and instruments that can be integrated into scholar’s workflows.’ (Evans 2008, *Harnessing digital and networking capabilities*, paragraphs 1 and 3)

## **Metadata**

‘It comes down to a view of valuing metadata as an integral part of the information infrastructure – not an adjunct to it. It is itself data and so with this perspective issues associated with the creation, management, migration and curation of scholarly content also apply to its metadata. With digital and networking technologies we are finally freed from the prison of the paper world where metadata is laboriously re-created at discrete moments in time, and forever remains disconnected. We no longer have to put up with minimal metadata but can instead look to have much richer, dynamic, descriptive metadata. It can be enriched on use – whether by machines (intelligent technologies, mining and associative analysis) or directly by humans (degrees of user annotation of metadata records). So I look forward to exciting times over the next ten years as the roles of scholars, other users, and information professionals in creating these kind of metadata management frameworks are negotiated.’ (Evans 2008, *Metadata* paragraph 1)

## **Citation and persistence**

‘As mentioned above scholarly information objects are inherently network in nature. Traditional scholarly information infrastructure is built on citation and persistence. In the paper world the citation of references in a scholarly work is a form of persistent identification. A citation is constructed to provide enough information to trace and locate the source so that the author’s assertions can be verified. In the digital world this can be made much more efficient with the addition of a direct link to the digital object itself. The persistence and stability of this referencing is vital if digital scholarly information is to endure through time. Conversely the benefits of having this persistent bridge between scholarly information objects opens up whole new ways of analysing, visualising and navigating scholarly spaces (Borgman 2007, p. 70).’

‘I would like to see the University seeking to realise the network nature of scholarly resources in its tools and services for management, discovery and access. The construction of exemplar digital archival citation frameworks is one area in which the University has the potential to show leadership given the strength of the University of Melbourne Archives and the relevance of its holdings to a number of research communities.’ (Evans 2008, *Citation and persistence*, paragraphs 1 and 3)

Key reference: Borgman, Christine (2007), *Scholarship in the Digital Age: Information, Infrastructure and the Internet*, MIT Press, Cambridge Massachusetts, ISBN 978-0-262-02619-2.

## Conclusion

This document is the first submission from the eScholarship Research Centre and it has attempted to provide a sense of the perspective from our experience of the University of Melbourne. Further documents will articulate some of the goals that we would like to see achieved in the short-term (2009), medium-term (2012) and in the longer term (2020). It is an interesting journey and one that we both watch and participate in with much scholarly interest.

Gavan McCarthy  
Director, eScholarship Research Centre  
9 May 2008