

---

## *Computing Arts 2001*

Digital Resources for Research in the Humanities, 26<sup>th</sup>-28<sup>th</sup> September 2001  
The University of Sydney

Thursday, 27<sup>th</sup> September 2001, 11am Parallel Sessions III  
A. Australian History Resources, Speaker Three

### **The Open Resource Scholarly Network: a new era for historians, archivists and technologists.**

Gavan McCarthy  
Australian Science and Technology Heritage Centre  
University of Melbourne

#### **Abstract:**

Since 1985 the staff of the Australian Science and Technology Heritage Centre (Austehc) have been collecting and disseminating information about the history of Australian science, technology and medicine including data about archival resources on the presumption that scholarly practice, including the creation of new knowledge, was based on free access to, and the citability of, existing knowledge. The advent of electronic network technologies has enabled us to realise our goals in ways that were not even dreams in the earlier environment. However, not all players in the scholarly information and publishing realms have responded in the same way. Despite these new technologies, which should be making resources much more readily accessible, many valuable resources are locked up (discoverable perhaps but uncitable) behind closed database walls or are available only on a user pays basis. In many cases these resources were previously available freely through research libraries. Austehc has spent the last few years developing database driven web publishing tools to support an open resource scholarly electronic network. These tools are being offered to the community at no cost, under the open source philosophy, if they are used for public good and education purposes. These tools, the "Online Heritage Resource Manager" (OHRM) and the "Web Academic Resource Publisher" (WARP) will be presented to this meeting.

#### **Introduction:**

The Australian Science and Technology Heritage Centre has been engaged in the collection, processing and dissemination of information about the history and heritage of Australian science technology and medicine since 1985. The Centre was originally a research project of Professor Rod Home in the Department of History and Philosophy of Science at the University of Melbourne. It was established as the Australian Science Archives Project and lived on a combination of soft money and semi-commercial projects until May 1999 when the Centre was formed to provide a better means by which the research, teaching and heritage functions of ASAP could be promoted and developed. Both ASAP and the Centre have been characterised by their early uptake of new technologies, in particular computer and communication technologies, to improve all phases of the history and heritage processes.

---

The aim of our research and heritage activities has been to support scholarly practice, particularly in the field of history of Australian science, technology and medicine. It was identified from early on that the tools, processes and practices developed would have applicability beyond our field and it has been a goal from the outset to produce generalised tools that could be used by heritage practitioners in a wide variety of settings. This was seen as one of the best means by which the knowledge we gained from our unusual and privileged position in the community could be translated into a form that would benefit others. By always seeking to generalise rather than particularise our tools we were forced to address the deeper issues facing archivists and heritage professionals, and, in particular, the challenges posed by the development of the Internet and the World Wide Web.

In using the Web to reach a wider user-base than was ever possible in the print world, we have sought to generalise and standardise the way content is presented. In other words, we looked for ways of presenting information derived from scholarly practice in a way that did not alienate the lay user, school students or the community at large. We wished to reach a wider community to build the generalised societal support that we saw as necessary for continuing development of the history and philosophy of science discipline. We were also becoming increasingly aware that much of the work being done in the history of Australian science, technology and medicine was being undertaken by enthusiasts, volunteers and others both outside and inside organisations, and that this knowledge was difficult to access.

We faced challenges on a number of fronts. We were seeking to:

- develop database-web process tools in a rapidly changing technological environment;
- publish materials and resources that would be long lasting;
- help the heritage and history communities evolve in concert with the changing technologies;
- ensure that the web technologies were developed to suit history and heritage needs;
- work within the heritage standards frameworks, many of which have only emerged within the last fifteen years.

In 1999 the Centre embraced the “open source” philosophy with regard to software and database tools and coined the term “open resource” to define the open scholarly citation network that we saw as a necessary requirement of a scholarly knowledge generation process based on electronic networks, in particular the Web. Very simply, what we wished to see was the process of scholarly citation that underpins academic discourse in the print environment fully implemented in the electronic networked environment. This seemed self-evident at the time but what happened in the “real” world was that commercial and self-interested forces came to control much electronic scholarly literature. The result being that much of this literature became inaccessible and uncitable. Even web resources published by scholars or explicitly for scholars often failed to meet the minimum requirements for building an electronic open resource scholarly network.

Moreover, an open resource scholarly network needs an information infrastructure, an infrastructure based on information and not on hardware and software, that will provide the contextual frameworks in which the resources can be meaningfully interrogated. Unfortunately much of the early, and indeed current web technology, does exactly the opposite - it de-contextualises resources. However, recent studies on the physics of the Web, analysing it as a complex network, has revealed that it is not a random network but what is described as a scale-free network which has much in common with natural complex systems and in particular human-friendly complex systems. Our experience has shown that there is much that can be done to direct the development of the Web so that it becomes a world of rich



contextualised interlinked spaces that could contribute much to support the development of human knowledge and understanding.

### **Citability, Coherence and Communicability**

An open resource scholarly network requires its component resources to meet the following criteria: citability, coherence and communicability. These are the key principles of Web object publication that will support the structuring of the Web (at least the 'heritage' space on the Web). Adherence to these principles will promote the development of the scale-free structure of the Web.

#### *Citability*

For a web object to be citable the user needs to know key information about the 'who, what, where and when' of the object origination. In the print world this was the imprint data and in the web world it is what has become known as "metadata".

Ideally, this information should be in both a human readable form and in a machine processable form. A standard for the machine processing form that has gained currency is the Dublin Core. The Dublin Core metadata set is used as the basis for the Australian government Web metadata standards, the Australian Government Locator Service or the AGLS. (For information on the Dublin Core Metadata Initiative see: <<http://dublincore.org/>> and the AGLS see: <[http://www.naa.gov.au/recordkeeping/gov\\_online/agls/summary.html](http://www.naa.gov.au/recordkeeping/gov_online/agls/summary.html)>).

An important aspect of citability is the "where". On the Web this is given by the URL or Uniform Resource Locator (See: <<http://www.w3.org/Addressing/URL/url-spec.html>>). Unfortunately despite the enormous power of the URL and its structure, not all Web publishing systems generate citable URLs. Although, in theory an eight character URL that utilised only the 36 alphanumeric characters could uniquely address over 2,821 billion individual Web objects, it is not uncommon to find URLs of the form:

```
<http://www.???.vic.gov.au/web/root/domino/cm_da/???ncor.nsf/frameset/???+Corporate?OpenDocument&[/4A25676D00279618/BCVIEW/AB5F9BE1DDCEAB074A25685D0016F9E0?OPENDOCUMENT]>
```

What is clear, as a necessary requirement of the human interface, is that URLs be both human readable and machine processable. For those interested in the problems of persistence of URLs go to <<http://purl.oclc.org/>>.

#### *Coherence*

Coherence refers to the hyperlink navigation from an object that will take you to the related web objects that the user must be able to access to understand the context, jurisdiction or framework in which the object has been published.

#### *Communicability*

Communicability refers to the ability of a web object to present its content in a standardised structured form that would enable the machine processing of elements of its content. It is taken for granted that

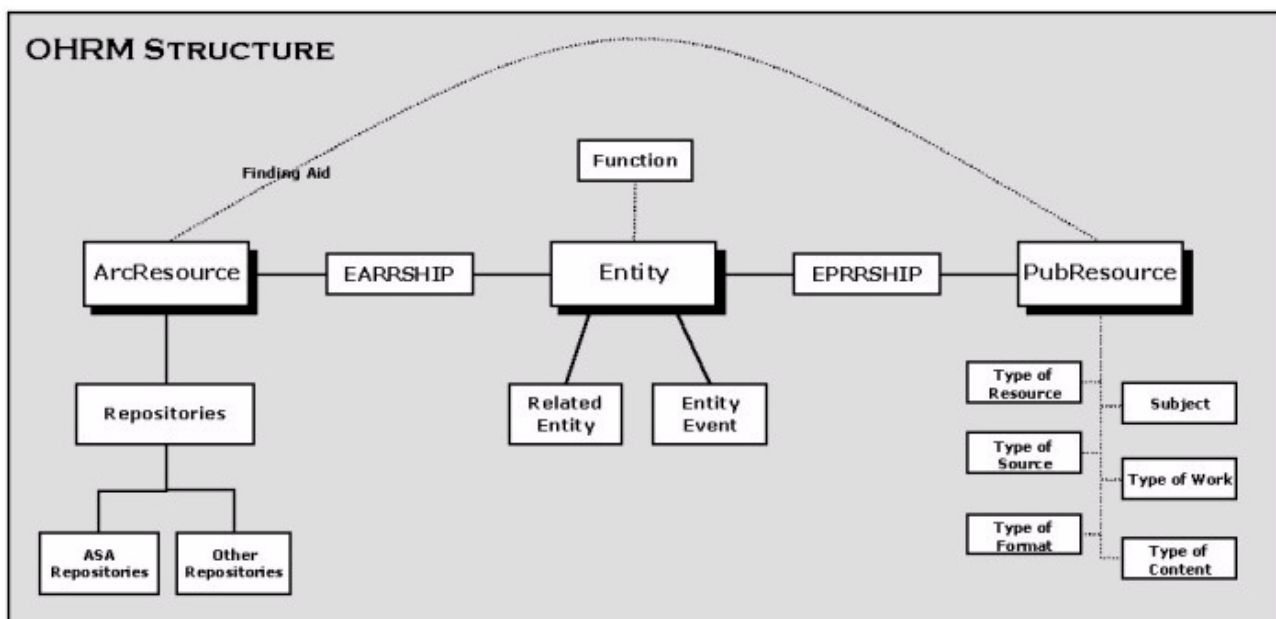
content should be expressed in a human readable form that assists the user to interpret the content. This is the context of visualisation. Hypertext Markup Language (html), the lingua franca of the Web, is geared to this end but not to the machine processing of the content. XML, or eXtensible Markup Language, has the capability to provide this service.

### Contextually bounded knowledge:

The Online Heritage Resource Manager or OHRM is a context based resource discovery and access system that links creators, archival and heritage resources and published materials within the one system.

The logical structure of the OHRM, as illustrated below, is based around:

- capturing key information about persons or corporate bodies in the Entity table;
- capturing key information about archival resources in the ArcResource table;
- capturing bibliographic citations of published resources, whether print, digital, online or offline in the PubResource table; and
- capturing the relationships of 'Entities' to other entities, archival and published resources in appropriate 'relationship' tables, i.e. RelatedEntity, Earrship and Eprship.



From this database, both static and dynamic HTML output can be generated to build a comprehensive web resource. More details about the OHRM can be found at <http://www.austehc.unimelb.edu.au/ohrm/>

### Web publishing that builds infrastructure:

The Web Academic Resource Publisher (WARP) is a database tool developed by the Australian Science and Technology Heritage Centre, to enable the scholarly web publication of reference texts. Promoting

more than just online reproduction of texts, the WARP facilitates the creation of a knowledge space which becomes a research tool from which new connections, insights and ideas can be discovered and explored.

Austehc has used the WARP to publish:-

*Technology in Australia 1788-1988*. The online edition of this bicentenary study by the Australian Academy of Technological Sciences and Engineering of the men, women and organisations involved in the development of technology in Australia. [<http://www.austehc.unimelb.edu.au/tia/titlepage.html>]

*Science and the Making of Victoria*. A Centenary of Federation project exploring the history and views of the Royal Society of Victoria since its inception in 1854 to the present day, and its role supporting science and technology in Victoria. [<http://www.austehc.unimelb.edu.au/smv/smv.html>]

*Victorian Patents and Patentees 1854 to 1904*. A pilot project in collaboration with the State Library of Victoria to explore the publication of Victorian patent applications from 1854 to 1904. [<http://patentsvictoria.net/>]

*Federation and Meteorology*. A Centenary of Federation publication on the emergence of Australian meteorology as a science and the formation of the Bureau of Meteorology in 1908, paralleling the story of Australian Federation. Compiled by the Australian Science and Technology Heritage Centre and the Bureau of Meteorology <http://www.austehc.unimelb.edu.au/fam/title.html>

The WARP is currently under development with the aim of making it available under license at no cost for non-commercial, heritage and public good purposes.

## **Conclusion:**

Austehc has uncovered a strong demand for the types of products outlined above. The OHRM, as a contextual framework builder and a key tool for creating scalable information infrastructure that utilises and enhances the “natural” tendency for humans to create scale-free complex networks that have strong fractal forms based on standardised semantic content. It has revealed itself to be adaptable to many different contexts and environments. The potential in the OHRM is substantial. The use of the OHRM in real situations with real data and with human operators with a range of foibles is robustly probing our ideas and concepts, and testing the structures and functions. This is challenging and immensely exciting.

The WARP has developed quickly and has moved from prototype to generalised tool with remarkable ease. We are currently working on the bridging mechanisms between the outputs from the OHRM and the WARP. Although XML has promised much it has been slow to deliver. We are looking to develop standardised outputs in XML that can be easily customised to capture individual identity and design without sacrificing functional and upgrade requirements.

Politically, we see the need for community financial support (i.e. support from government) to enable the building and maintenance of the open resource scholarly network. It is a bit like a virtual national park - it is a place for all people to visit, a place in which knowledge and experience can be shared, a place where all are welcome, a place where our heritage is conserved and transmitted.



**-oOo-**