



# The Life of Information

**A one-day symposium on  
the design and use of  
online dictionaries,  
encyclopedias &  
collections**

**Twitter** — use hash tag #lois2010

**Internet** — username: a326228  
password: lois2010

Friday 24 September 2010, Sir Roland Wilson Building, ANU

# >Program

8.30-9.00am	REGISTRATION
SESSION 1	<b>Lives Online</b> CHAIR: Nicholas Brown, Senior Fellow, School of History, ANU
9.00am	<b>WELCOME</b> Professor Robin Stanton, Pro Vice-Chancellor (E-Strategies), ANU Paul Arthur, Deputy Director, National Centre of Biography, ANU
9.25am	<b>Stephen Due</b> <i>Australian Medical Pioneers Index: Research and Collaboration in Digital Lives</i>
9.50am	<b>Janet McCalman, Sandra Silcot &amp; Len Smith</b> <i>Founders &amp; Survivors: Australian Life Courses in Historical Context, 1803-1920</i>
10.15-10.45am	MORNING TEA
SESSION 2	<b>Stories in the Archive</b> CHAIR: Debjani Ganguly, Head, Humanities Research Centre, ANU
10.45am	<b>Zoë D'Arcy</b> <i>The Art of Online Story-Telling: Exposing the National Archives' Collection</i>
11.10am	<b>Cassandra Pybus</b> <i>Recovered Lives: Constructing a Digital Repository for Enslaved Africans</i>
11.35am	<b>Katherine Bode</b> <i>Speed Reading in the Digital Age: The Case of Women's Writing</i>
12.00-1.00pm	LUNCH
SESSION 3	<b>FRIDAY FORUM: Flexibly Digital</b> FACILITATOR: Paul Arthur, Deputy Director, National Centre of Biography, ANU
THIS FRIDAY FORUM IS JOINTLY HOSTED BY THE RESEARCH SCHOOL OF HUMANITIES & THE ARTS, CONVENED BY MARY HUTCHISON & ALISON WAIN	
1.00pm	Introduction to the Friday Forum
1.15pm	Presenters: <b>Donald Hobern, Kerry Taylor &amp; Basil Dewhurst</b>
2.00pm	Discussion
2.30-3.00pm	AFTERNOON TEA
SESSION 4	<b>Knowledge Networks</b> CHAIR: Paul Turnbull, Professor of eHistory, University of Queensland
3.00pm	<b>Ian Johnson</b> <i>Structure after the Fact: Abstract Databases and Digital Encyclopaedias</i>
3.25pm	<b>Ross Coleman, Emma Grahame, Steven Hayes &amp; Stewart Wallace</b> <i>Dictionary of Sydney</i>
3.50pm	<b>Tim Sherratt</b> <i>Liberating Lives: Invisible Australians and Biographical Networks</i>
4.15pm	SYMPOSIUM CLOSE
From 4.30pm	DRINKS FOR PURCHASE AT UNIVERSITY HOUSE GRADUATES ROOM (DOWNSTAIRS)

Stephen Due

*Australian Medical Pioneers Index: Research and Collaboration in Digital Lives*

This paper discusses the design and use of the *Australian Medical Pioneers Index*, a national biographical database of colonial doctors from settlement to 1875. The project is distinguished from more conventional, narrative-based approaches to collective biography by its highly-structured database format. The database was designed specifically to store biographical information so that it could be used in historical research. In this respect it might be compared to a system of electronic medical records, each of which serves as the history of an individual patient, but which collectively can be used to study patterns of disease in a community.

At the present time the data set is not fully mature, but preliminary analyses of the data have looked successfully at the qualifications of doctors emigrating to Australia, the range of government posts available to them in the colonies, and suicide in the profession. Meanwhile the database design has proved to be fruitful as a biographical tool in its own right, and the paper considers its advantages and disadvantages in this respect. One advantage has been the opening up of numerous opportunities for collaboration with other projects that have a significant, overlapping biographical component. To date these have included projects in such diverse fields as family history, military history and local history. The paper concludes with a discussion of the values and objectives of the project in the context of a particular tradition in medical biography (that of professional rolls) and a consideration of its future prospects.

Janet McCalman, Sandra Silcot & Len Smith

*Founders & Survivors: Australian Life Courses in Historical Context, 1803-1920*

*Founders & Survivors* is a partnership between historians, genealogists, demographers and population health researchers. It seeks to record and study the founding population of 73,000 men, women and children who were transported to Tasmania, and from that base build a longitudinal study of their descendants concluding with the military service records in the National Archives of Australia. The project is an exercise in interdisciplinary prosopography that will result in the world's first cradle-to-grave datasets that can be linked intergenerationally, and will be mined by demographers, epidemiologists, historians, social scientists and family historians. It is also an exercise in advanced digital humanities, employing TEI (Text Encoding Initiative), which is an established XML encoding standard designed for working with manuscripts. The data is organised in cascades of linked records using a program devised by Dr John Bass AM for the medical records of Western Australia.

Research team: Assoc Prof Hamish Maxwell-Stewart, University of Tasmania (lead CI); Prof Janet McCalman, University of Melbourne; Dr Rebecca Kippen, University of Melbourne; Assoc Prof Ralph Shlomowitz, Flinders University; Assoc Prof Shyamali Dharmage, University of Melbourne; Assoc Prof Alison Venn, Menzies Research Institute, University of Tasmania; Dr John Bass, Menzies Research Institute, University of Tasmania; Gavan McCarthy, University of Melbourne; Dr Len Smith, Australian National University; Dr Peter Gunn, Flinders University; Dr Alison Alexander, University of Tasmania; Dr Deb Oxley, Oxford University; Assoc Prof David Meredith, University of New South Wales; Dr James Bradley, University of Melbourne; Sandra Silcot, University of Melbourne; Claudine Chionh, University of Melbourne; Dr Robin Petterd, Sprout Labs, Hobart. International partners: Prof Richard Steckel, University of Ohio, Columbus; Prof Kris Inwood, Guelph; Prof John Cranfield, Guelph; Prof LH Lumey, Mailman School of Public Health, Columbia.

---

Zoë D'Arcy

*The Art of Online Story-Telling: Exposing the National Archives' Collection*

The online audience is an important one for the National Archives. For many years now the Archives has been reaching a much bigger audience online than could possibly be reached through its Reading Rooms, exhibitions or public programs. However, when people physically walk into the Archives their experience can be shaped, monitored and evaluated to ensure that their visit is a success. The Archives can weave stories for them from the collection through exhibitions and other programs, or work with them in the Reading Room to help them research their own stories. The same degree of control is not true of the Archives' online presence. What sort of experiences are online audiences having of the Archives, given the aging structure of both the Archives website and its collection database RecordSearch? What sort of experience could they potentially be having? What stories could potentially be discovered and told from the Archives' collection?

In 2008 the National Archives, in partnership with the Department of Veterans' Affairs, created the website *Mapping Our Anzacs*, which gave online audiences a whole new way of experiencing the digitised WWI files of Australia's servicemen and women. The richness of the site struck an instant chord with online audiences, who could use the site to both find and tell their own family or community's stories. Its creative place-based interface, in combination with the way in which it made use of user contributions, gave the Archives a whole new way of viewing how its records could be accessed online. This paper will look at how the Archives, using *Mapping Our Anzacs* as a starting point, is exploring how to combine its two most powerful assets – an online collection and an online audience.

---

Cassandra Pybus

*Recovered Lives: Constructing a Digital Repository for Enslaved Africans*

Historians continue to discuss enslaved Africans the way their owners did, as an undifferentiated mass. The defensive response has been that slaves were largely illiterate, so, with few exceptions, they left no individual accounts and as they carried no surnames we cannot trace individuals in archival records. This is a proposition rightly rejected by the Pulitzer prize winner, black historian Annette Gordon-Reed, who admonishes historians for falling back on homogenised accounts of the experience of enslaved people: 'The erasure of individual black lives – indeed the assumption that the concept of individual as opposed to group identity is meaningless', she writes, 'saves historians the trouble of having to discover and analyse the individual experience'.

Working on the principal that enslaved African Americans were individuals with complex life experiences, I have developed the website *Black Loyalist* as a repository of data to provide as much biographical information as can be found for the individual slaves who ran away to join the British during the American Revolution. Luckily the British army kept a meticulous list, known as the Book of Negroes, which records of the personal details of some 3000 runaways evacuated as free people in 1783, organised by ship and destination, with each person given a name, age, brief description, owner's name, date of absconding, and geographic location. With intensive archival research, many of these people can be located in a historical landscape, which can tell us a lot about their lives.

In the pilot project, about 600 people from Virginia who went to Nova Scotia have been traced backwards and forwards in the historical record. Linking this data digitally has allowed me to construct life trajectories, kin relationships, religious affiliations and more, that completely confound the received wisdom about the nature of slave experience in the eighteenth century.

---

---

Katherine Bode

*Speed Reading in the Digital Age: The Case of Women's Writing*

The growth, over the past two decades, in digital humanities infrastructure has dramatically expanded the quantity and range of information available to humanities scholars. However, as Jerome McGann asserts, 'the general field of humanities education and scholarship will not take the use of digital technology seriously until one demonstrates how its tools improve the ways we explore and explain aesthetic works – until, that is, they expand our interpretational procedures' (*Radiant Textuality*, xii). Proving the relevance and interpretive potential of electronic resources is, in other words, a foundational and defining challenge of digital scholarship.

My project takes up this challenge in relation to *AustLit: The Australian Literature Resource*. Extracting and analysing data from this electronic archive has enabled me to identify and explore hitherto unrecognised trends in Australian publishing, authorship and criticism, while optimising the utility of an existing digital resource. This paper will demonstrate the exploratory and interpretative potential of such research in relation to gender trends in Australian literature. My analysis complicates the routine attribution of growth in the proportion of women writers to feminism, while suggesting the role of other factors – such as genre, publishing industry dynamics and literary criticism – in shaping gender trends in authorship.

---

FRIDAY FORUM

*Flexibly Digital*

The thick, descriptive data of humanities research and the strict categorisation of digital databases are not natural partners. Current developments in data representation are breaking down these barriers. This forum discusses the theory and practice of flexible digital search techniques.

Donald Hobern will talk about the *Atlas of Living Australia* (ALA), a government-funded initiative to integrate information on Australian biodiversity. The ALA exemplifies the way new data representations can bring diverse sources of information together – from herbaria, to literature, to observations from the public.

Kerry Taylor will discuss the largest knowledge representation project ever attempted, the Semantic Web, which uses data representations that have a heritage in natural language processing. These make data accessible in many different ways, avoiding the rigid structure of the traditional database.

Basil Dewhurst will use examples from the National Library of Australia's Trove People and Organisations program and the Australian National Data Service (ANDS) to discuss the creation, maintenance and sharing of data to support research.

---

---

Ian Johnson

*Structure after the Fact: Abstract Databases and Digital Encyclopaedias*

The *Dictionary of Sydney* is built on top of a generic web database (Heurist) designed from the ground up for humanities research data. Heurist uses an abstract data model which can accommodate any type of physical or conceptual entity (building, map, document, person, event, role, relationships, annotations etc.) without any modification of the underlying database structure or effect on existing data.

The independence between database structure and the domain modelled confers the flexibility required by open-ended humanities projects and encourages the granular recording of information (for example, birth, marriage and death as individual fact records rather than as fixed calendar attributes of individuals). The remixing possibilities of such granular data allow decisions about delivery formats to be taken after-the-fact, allowing data to be repurposed for web sites, data feeds, maps and mobile applications etc.

I will explain the simple concepts behind an abstract model of humanities data, illustrated by comparing Heurist with the conventional database structure developed for the *Digital Harlem* project, and demonstrate the advantages of an abstract model approach. I will focus particularly on the ability to create a typed, directional relationship between any two entities in the database, and to annotate and timestamp this relationship. This function has been used to great effect in the *Dictionary of Sydney* as a means of enriching the entries with links to related material, as a navigation device within the *Dictionary*, and for the dynamic construction of maps and timelines.

The *Dictionary* uses a generic XML output function styled through Cocoon and XSLT to generate the public website. I will finish by discussing future development directions including a version for mobiles, and the scheduled Open Source release of Heurist version 3. This version will streamline on-demand creation and customisation of new databases.

---

Ross Coleman, Emma Grahame, Steven Hayes & Stewart Wallace  
*Dictionary of Sydney*

The *Dictionary of Sydney* is producing a new kind of history of Sydney: online, growing and changing, covering every aspect of human life in this place. The website is the first window into the *Dictionary* but, over time, it will also be accessible through other technologies, including mobile delivery, print-on-demand and others yet to be developed.

Geographically, the *Dictionary of Sydney* includes the whole Sydney basin and spans the years from the earliest human habitation to the present. The project welcomes all kinds of history – engineering history, social and cultural history, economic history and so on. We are also interested in historical contributions from neighbouring disciplines such as archaeology, sociology, literary studies, historical geography and cultural studies.

The *Dictionary* contains new historical writing enriched with other resources – images, maps, sound, music, oral history, film, documents and pointers to important objects in Sydney collections. Contributors include academics, professional historians, heritage specialists, local studies experts, genealogists, enthusiasts, volunteers and readers.

This session reports on progress to date, and plans for the future.

---

---

Tim Sherratt

*Liberating Lives: Invisible Australians and Biographical Networks*

‘What changes’, asks Tim Hitchcock, ‘when we examine the world through the collected fragments of knowledge that we can recover about a single person, reorganised as a biographical narrative, rather than as part of an archival system?’ Resources such as *People Australia* and the *Australian Dictionary of Biography* provide us with a solid framework for using and sharing biographical information. But what about those people whose lives were not so grand, whose achievements were not so notable, whose remnants are not so easily found? Elsewhere online, family history is flourishing, fuelled by the growing range of genealogical resources. Through name indexes, cemetery registers, case files and more, the scattered fragments of individual, ordinary lives are being liberated from the archives. How do we link these fragments? How do we find space within our biographical infrastructure for the exploration of ordinary lives?

Kate Bagnall and I are embarking on a project to recover ‘invisible Australians’ from the bureaucratic remains of the White Australia Policy. We plan to crowdsource the extraction of data from the many thousands of forms and documents used to monitor the movements of non-white Australians. But we want to do more. We want to work with community groups and family historians to enable them to expose their own data in re-usable ways. We want to link down to a lonely headstone in a small country graveyard, and publish up to *People Australia* and beyond. As well as databases, we want to build capacities, tools, skills and collaborations. We want to create an inclusive, participatory framework within which we can seek an answer to Tim Hitchcock’s question. What will the world look like through this network of liberated lives?

---

Paul Arthur is Deputy Director of the National Centre of Biography and Deputy General Editor of the *Australian Dictionary of Biography*, Australian National University. He has published widely on the history of technology, media, travel and empire, including *Virtual Voyages* (2010) and *History and New Media* (forthcoming in 2011). An advocate for digital humanities and e-research, he is series editor of Anthem Scholarship in the Digital Age (Anthem Press, London & New York).

Katherine Bode is Lecturer in English, University of Tasmania. From 2011 she will join the Research School of Humanities and the Arts at ANU as Senior Lecturer in Digital Humanities. Her current project investigates the critical potential of quantitative analysis to uncover new knowledge about Australian literary and publishing history, and the history of the novel more broadly. Her latest monograph, *Reading by Numbers: Recalibrating the Literary Field*, will be published in 2011.

Ross Coleman is Director of Sydney eScholarship at the University of Sydney Library. He has extensive experience in the creation and management of digital library collections, including the SETIS (Scholarly Electronic Text and Image Service) primary source collections in Australian Studies. He was instrumental in re-establishing Sydney University Press as an electronic / print publisher, and the establishment of Sydney eScholarship, providing digital archiving and publishing services to the University.

Zoë D'Arcy is the Director of Communications and Programs at the National Archives of Australia. This section encompasses programs such as exhibitions, publications, websites and events, as well as media and marketing. Her focus is on establishing the relevance of the National Archives and its collection to a wide national audience.

Basil Dewhurst is Project Manager, Australian Research Data Commons (ARDC) Party Infrastructure Project, a project designed to enable the improved discovery of, and better management and curation of, research outputs and data in Australia. He has a longstanding interest in metadata standards and data interchange and prior to his current role he managed a number of the National Library's discovery services: Australian Research Online, Music Australia, People Australia and Picture Australia and was the Project Manager of the People Australia Project. He is active on the international EAC Working Group, the body charged with developing the EAC-CPF standard for describing people organisations and families (released in February 2010), and the Standards Australia IT019 committee. Prior to working at the National Library he was Manager, Image Services, at the Powerhouse Museum and Technical Developer for Australian Museums OnLine (AMOL).

Stephen Due is Chief Librarian at Barwon Health, Victoria, and Editor of the *Australian Medical Pioneers Index*. He has been working with digital biographical resources since 1999 and has been editor of AMPI since 2001. His particular areas of expertise are colonial doctors and Australian doctors at war. Apart from the AMPI website, his publications include two bibliographies and a number of journal articles.

Emma Grahame is Editorial Coordinator of the *Dictionary of Sydney* project. She is employed by the University of Sydney under the Australian Research Council and joined the *Dictionary* in May 2007. Emma has extensive experience in print and digital editing as well as in historical research.

Steven Hayes is Technical Coordinator of the *Dictionary of Sydney* project and is Project and Business Development Manager at the Archaeological Computing Lab at the University of Sydney. The Lab is the base for technical research on the project. Steven has a background in the practical design and development of websites and other projects focused on furthering environmental and social sustainability.

Donald Hobern is the Director for the *Atlas of Living Australia* (ALA), a project which aims to deliver information systems to integrate information about all Australian species of plants, animals and microorganisms to support research, education and decision-making. The ALA will launch its first public web tools in late 2010. He is also Chair of the Biodiversity Information Standards (Taxonomic Database Working Group – TDWG), the international organisation responsible for development of standards for exchange of biodiversity data.

Ian Johnson is Director of the Archaeological Computing Laboratory and Deputy Director of the Digital Innovation Unit, University of Sydney. He has been teaching digital methods in archaeology and developing tools for archaeology and for the digital humanities since the 1980s, with a particular emphasis on collaborative databases, GIS, mapping and the temporal dimension. From 1987–2000 he developed the computerisation of the NSW Aboriginal Sites Register and from 1998–2003 the technical infrastructure for the Electronic Cultural Atlas Initiative. His major development projects include Minark (1980–1987; a database management system for archaeology), TimeMap (1997–2003; time-enabled distributed-data mapping), Heurist (2005 onwards; collaborative bibliographic system and database) and FieldHelper (2004 onwards; management and metadata creation for fieldwork data). These software tools have provided the technical infrastructure for projects such as *Dictionary of Sydney*, Macquarie Library's *MacquarieNet* online encyclopaedia, *Macquarie Atlas of Indigenous Australia* and ABC Innovation's *Gallipoli: The First Day*.

Janet McCalman holds joint appointments in History and Philosophy of Science in the Faculty of Arts and in the Centre for Health and Society in the Faculty of Medicine, Dentistry and Health Sciences, at the University of Melbourne. She is a Fellow of both the Academy of the Humanities and the Academy of the Social Sciences. She has published two histories of Australian life and politics, *Struggletown* (1984, 1998) and *Journeyings* (1993). The social history of women's health, *Sex and Suffering* (Melbourne University Press 1998), was also published in the United States by Johns Hopkins University Press. Her current research interests are in the social history of health and disease, life course history, the family, and ecological history.

Cassandra Pybus is an Australian Research Council Professorial Fellow in History at the University of Sydney. She has interests that span Australian social history, colonial history in North America, South East Asia, Africa and Australia, slavery and the history of labour. Her most recent books are *Epic Journeys of Freedom: Runaway Slaves of the American Revolution and Their Global Quest for Liberty* (2006), *Black Founders: The Unknown Story of Australia's First Black Settlers* (2006), and as co-editor, *Many Middle Passages: Forced Migration and the Making of the Modern World* (2007).

Tim Sherratt is a digital historian, web developer and cultural data hacker who has been developing online resources relating to archives and history since 1993. He has written on weather, progress and the atomic age, and has developed resources including *Bright Sparcs* and *Mapping our Anzacs*. Currently employed by the Australian National Data Service and the National Museum of Australia, he is also an Adjunct Associate Professor in the Digital Design and Media Arts Research Cluster, University of Canberra.

Sandra Silcot is a systems developer who is also an honours graduate in history and social theory from the University of Melbourne. She has worked in digital humanities and enterprise systems at the University since the early 1990s.

Len Smith of the Australian Demographic and Social Research Institute at the Australian National University is a demographer on the *Founders & Survivors* project and a key advisor on data management strategies.

Kerry Taylor is a principal research scientist in CSIRO and teaches on the Semantic Web in the Faculty of Engineering and Information Technology, ANU. Since 1995 she has worked in CSIRO on data and system integration technologies, predominantly in areas applied to natural resource management but also in bioinformatics, business enterprise, health and defence.

Stewart Wallace is Project Manager of the *Dictionary of Sydney*. He is employed by the University of Sydney under the Australian Research Council and commenced as Project Manager in May 2006. For the previous 10 years, Stewart worked at the City of Sydney in a variety of IT-related roles, most recently as IT Development Manager.



**Thank you for attending the Life of Information symposium.  
To be notified of future events, please contact the convenor:**

Dr Paul Arthur  
Deputy Director, National Centre of Biography  
Deputy General Editor, Australian Dictionary of Biography  
School of History, Research School of Social Sciences  
H. C. Coombs Building, No 9  
The Australian National University  
Canberra ACT 0200

Email: [paularthur@anu.edu.au](mailto:paularthur@anu.edu.au)  
Web: [paularthur.com](http://paularthur.com)  
Twitter: [@pwlarthur](https://twitter.com/pwlarthur)