

DIGITAL HUMANITIES AUSTRALASIA 2014:

Expanding Horizons

17-21 March, 2014, The University of Western Australia

The aim of DHA 2014 is to advance digital methods, tools and projects within humanities research and develop new critical perspectives. The conference will provide a supportive, interdisciplinary environment to explore and share new and advanced research within the digital humanities.

Hosts and Organisation Committees

The conference is being organised via the iVEC@UWA Facility.



The Local Organisation Committee consists of:

Paul Arthur, The University of Western Sydney

John Hartley, Curtin University

Charise Baker, The University of Western Australia

Jo Hawkins, The University of Western Australia

Paul Bourke, The University of Western Australia

Philip Mead, The University of Western Australia

Toby Burrows, The University of Western Australia

Renee Newman-Storen, Edith Cowan University

Jenni Harrison, The University of Western Australia

Academic Programme Committee consists of:

Paul Arthur, The University of Western Sydney

Ian Johnson, University of Sydney

Linda Barwick, University of Sydney

Nagasaki Kiyonori, The University of Tokyo, Japan

Craig Bellamy, University of Melbourne

Gavan McCarthy, University of Melbourne

Katherine Bode, Australian National University

Sydney Shep, Victoria University Wellington, NZ

Erik Champion, Curtin University

Tim Sherratt, National Museum of Australia

Arianna Ciula, EADH, London, UK

Harold Short, Kings College London, UK

Hugh Craig, Newcastle University

James Smithies, University of Canterbury, NZ

Jenni Harrison, University of Western Australia

Nick Thieberger, The University of Melbourne

Brett D Hirsch, University of Western Australia

Deb Verhoeven, Deakin University

Jane Hunter, University of Queensland

This event would not have been possible without the support of the following sponsors:

- iVEC@UWA
- Curtin University
- Edith Cowan University
- The University of Western Australia
- Perth Convention Centre, and
- The Australian Literature Westerly Centre.

Platinum Sponsors



Gold Sponsors



Silver Sponsors



**Australian Literature,
Westerly Centre, UWA**

Table of Contents

Conference Information.....	4
Outline	8
Programme.....	9
Wednesday, 19 March 2014 (Conference Day 1 - summary).....	10
Keynote - Professor Neil Fraistat	10
Thursday, 20 March 2014 (Conference Day 2 - summary).....	11
Keynote - Dr Sarah Kenderdine	11
Friday, 21 March 2014 (Conference Day 3 - summary)	12
Keynote – Dr Anthony Beavers.....	12
Pre-conference workshop details.....	13
Main Conference - Wednesday, 19th March 2014.....	21
1.1: Language & Text I	21
1.2: Print Cultures	25
1.3: Online Engagement & Communities.....	27
2.1: Language & Text II	30
2.2: Archives & Collections I	34
2.3: BOF: Digital Humanities and the Tyranny of Distance	37
3.3: BOF: The Altruistic Crowd – The Ethics of Using Social Data in Humanist Research?	45
Main Conference - Thursday, 20th March 2014.....	46
4.1: Digital History & Virtual Heritage I.....	46
4.2: Network Analysis & Discovery	50
4.3: BOF: Building DH Community on the Regional Scale - Notes Towards a Multimodal Regional Digital Cultures Collaboratory.....	54
5.1: Digital History & Virtual Heritage II.....	55
5.2: Infrastructure & Collaboration	58
5.3: BOF: All of the Records, All of the Time.....	61
6.1: Digital History & Virtual Heritage III.....	62
6.2: Geo-Spatial Methods	64
6.3: BOF: Digital Humanities Afield.....	67
Main Conference - Friday, 21st March 2014.....	69
7.1: Visualisation I.....	69
7.3: BOF: Debate - That Literary Studies Needs More Graphs, Maps, and Trees	72
8.1: Visualisation II.....	73
8.2: Literature & Literacies	76
9.1: Visualisation II.....	80
9.2: Digital Editions & Editing	81
9.3: Digital Evidence	84

Conference Information

INFORMATION ONLINE

PDF versions of conference brochure are available at <http://dha2014.org/wp-content/uploads/2014/02/DHA2014-brochure.pdf>

VENUE

Conference sessions will be held at the University Club, The University of Western Australia, Crawley. Venues are open throughout the conference so please be sure to keep your valuable items with you.

MAP

The location of the University Club (shown in white) is provided below.

<http://www.universityclubconferences.com.au/template.asp?CID=4>



The University Club is situated just eight minutes drive from the Perth CBD on the banks of the Swan River, The University Club is located on campus at The University of Western Australia on Hackett Drive in Crawley (Hackett Entrance 1).

REGISTRATION

The conference registration desk is in the terrace foyer of the University Club, located off the rear garden. It will be open from 8am on 19th March and 8:30am thereafter.

No Late registrations will be taken

Please register for the conference no later than 9th March 2014, midnight, WST.

All registrations must be made online via the following link:

<http://payments.weboffice.uwa.edu.au/mech/DHA2014>

REGISTRATION COST PER PERSON

aaDH/ADHO Members Full Registration:	\$535.00
Full Registration (non aaDH/ADHO Members):	\$685.00
Student Full Registration	\$200.00
Day Delegate Registration	\$200.00

PARKING

If you are driving to the conference, plan to arrive early. Visitor parking at the University of Western Australia is clearly signposted but is limited. Come prepared with adequate coin change.

CONFERENCE RECEPTIONS

There will be several conference receptions:

1. Tuesday 18th March (Edith Cowan University) - Musical evening reception

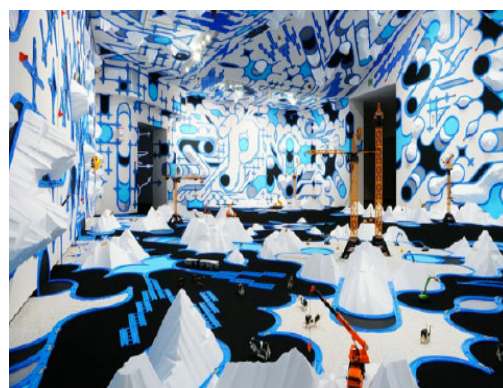
An evening reception (sundowner) is being held at the Spectrum Gallery, Mt Lawley Campus, ECU. This event will include a recital from WA Academy of Performing Arts "Decibel" and the gallery will also be showcasing wall drawings by the artist Marie Haass and others.



This will follow the free Masterclass, hewing hosted by Anthony Beavers.

2. Wednesday 19th March (Curtin University) – Visualisation evening reception

Professor Tim Dolin, Humanities Dean Research and Graduate Studies, welcomes you to a Visualisation Evening and Poster reception (sundowner) at the Curtin HIVE and John Curtin Gallery. Dr Hutchison will present virtual environments from the Sydney Kormoran Project and Professor Stelarc will present a performance art piece integrating immersive visualisation, avatars and body tracking. Transport will be provided.



Paramodelic - Graffiti, 2012, PARAMODEL, Tokyo Station Gallery
Photo: paramode

Both of these receptions are available even to those not attending the conference and the cost per person is \$25.00 per person; **this cost covers both events.**

3. Thursday 20 March 2014 (UWA) - Alexander McCall Smith – Minor Adjustment Beauty Salon

Book reading with Alexander McCall Smith – provided by UWA Extension at UWA's Octagon Theatre (Thursday 20th March) – **PAYABLE to UWA Extension only.**

The talk will run for one hour at the Octagon Theatre (at UWA) with book sales following.

An hour in the company of Alexander McCall Smith is an hour of sheer delight. This master story-teller will take you on a journey between Scotland and his characters from Scotland Street and the Isabel Dalhousie novels, Botswana and the much-loved Precious Ramotswe, and the trains of the East Coast of the United Kingdom and the Australian Outback – with a great deal of good cheer along the way. Modern ideas get tangled up with traditional ones in the latest intriguing Botswana book from the bestselling author, Alexander McCall Smith. Join us at UWA to hear the author talk about Mma Ramotswe's fourteenth adventure, The Minor Adjustment Beauty Salon. He will also introduce you to his much anticipated new, standalone novel – The Forever Girl.

Tickets are modestly priced at: \$39.00 (adult), \$30.00 (concession) & \$20.00 (student)

More details are available using the link <http://www.extension.uwa.edu.au/course/CCVW001>

CATERING

Morning tea will be served each morning in the Terrace Foyer of the University Club. The conference registration fee also lunches and afternoon teas on all days. There will be plenty of vegetarian and non-vegetarian options. Anyone with food allergies must indicate this to the conference organisers **in advance** of the event. Water will be provided at all venues. Please note that food or drink (except bottled water) is not permitted in the conference venues.

LOCAL CAFES

There are local cafes on Broadway (5 minutes walk from UWA Campus) as well as the Uniclub café and various Guild cafes located on the UWA campus. Matilda bay café is also only a short walk away.

WIRELESS INTERNET

An unsecured network is available inside the University Club. For information on how to connect to EDUroam at UWA, see the following information: <http://www.is.uwa.edu.au/it-help/access/wireless/eduroam>

INFORMATION FOR ORAL PRESENTERS

Please bring your presentation on a standard USB stick or your own PC / Mac. There is no need to send presentations in advance. Plan to load and test your presentation on the computer in your venue during the break before your presentation. If you plan to have a dynamic presentation, relying on the

internet, please sure to have a static version of your presentation (i.e. screenshots of web pages or dynamic content) as backup.

Sessions typically have a mix of long and short papers, and you will need to be aware of what type of paper you are presenting. Long papers are 20 minutes, with 10 minutes for questions. Short papers are 10 minutes, with 5 minutes for questions. Session Chairs will have flashcards to indicate the time remaining for presenters, and should be strict about keeping to the schedule. No additional software or plugins can be loaded onto conference computers. Podium, lapel or hand-held microphones are available in all venues.

INFORMATION FOR POSTER PRESENTERS

There is no standard prescribed poster size. Posters must be supplied in electronic format (either .pdf or .ppt format). Authors should contact the organisers for details.

INFORMATION FOR SESSION CHAIRS

Many sessions have a mix of long and short papers, and session Chairs are responsible for keeping presenters to time. Long papers are 20 minutes, with 10 minutes for questions. Short papers are 10 minutes, with 5 minutes for questions. Flashcards with '10 minutes', '5 minutes' and '2 minutes' notifications are provided in all venues, for use by Chairs.

TWITTER

Join a bigger conversation and use the hash tag #dha2014 when you tweet about this event.

QUESTIONS?

For any questions or administrative issues, please contact dha2014@ivec.org

Outline

The outline for the conference week will be as follows:

Mon 17th March	Workshops
Tues 18th March	Workshops Masterclass with Anthony Beavers (Computational philosophy and the moral implications of automated decision making, at ECU Mt Lawley) Musical Reception (evening sundowner, Spectrum Gallery, ECU Mt Lawley)
Wed 19th March	Conference (full day) Visualisation and Poster Reception (evening sundowner, John Curtin Galley, Curtin University)
Thurs 20th March	Conference (full day) Book reading with Alexander McCall Smith (Octagon theatre, UWA)
Fri 21st March	Conference (full day)

Please note this programme is subject to change without notice.

Programme

Monday, 17 March 2014 (Workshops)

Time	Title	Presenter	Location
9:00 – 17:00	Workshop– Computation Based Discovery in Literary Language	Hugh Craig, Jack Elliott & Brett D. Hirsch	CLT, Ground Floor Physics Building, UWA Crawley Campus
14:00 – 17:00	Workshop – Create beautiful scholarly maps with TileMill	David Flanders & Steve Bennett	CLT, Ground Floor Physics Building, UWA Crawley Campus
14:00 – 17:00	Workshop – Boosting your research profile through your research data	Adrian Burton	CLT, Ground Floor Physics Building, UWA Crawley Campus

Location for all workshops on 17th March is as follows:

- **Centre for Learning Technology** - Ground floor Physics building, UWA Crawley campus

Tuesday, 18 March 2014 (Workshops)

Time	Title	Presenter	Location
9:00 – 12:00	Workshop – #altac Track: Strategies to imagine and build alternative academic careers'	Jason Ensor & Jo Hawkins	CLT, Ground Floor Physics Building, UWA Crawley Campus
9:00 – 12:00	Workshop – Federated Archaeological information Management Systems	Brian Ballsun-Stanton & Ian Johnson	CLT, Ground Floor Physics Building, UWA Crawley Campus
9:00 – 12:00	Workshop – Automated 3D Model Reconstruction from Photographs	Paul Bourke	CLT, Ground Floor Physics Building, UWA Crawley Campus
9:00 – 12:00	Workshop – A taste of HuNI	Toby Burrows & Alex Hawker	Read Library Training Room , UWA Crawley Campus
14:00 – 17:00	Masterclass – Computational philosophy and the moral implications of automated decision making	Anthony Beavers	iVEC@ECU , Building 13, ECU Mt Lawley Campus
14:00 – 17:00	Workshop – A taste of HuNI (Repeat)	Toby Burrows & Alex Hawker	Read Library Training Room , UWA Crawley Campus

Location for workshops on 18th March is as follows:

- **Centre for Learning Technology** - Ground floor Physics building, UWA Crawley campus
- **Read Library Training Room** – Reid Library, UWA Crawley Campus
- **iVEC@ECU** - Building 13, Edith Cowan University Mt Lawley Campus

Wednesday, 19 March 2014 (Conference Day 1 - summary)

8:30am	Registration		
-	Location: University Club Ground Floor Terrace		
9:00am			
9:00am	Opening		
-	Location: University Club Auditorium		
9:15am	Professor Paul Johnson, Vice Chancellor of The University of Western Australia		
9:15am	Keynote - Professor Neil Fraistat		
-	Location: University Club Auditorium,		
10:30am	Chair: Professor Paul Arthur, University of Western Sydney / President, Australasian Association for Digital Humanities		
10:30am	Morning Tea		
-	Location: Ground Floor Terrace /Lower Colonnade		
11:00am			
11:00am	1.1: Language & Text I	1.2: Print Cultures	1.3: Online Engagement & Communities
-	Location: University Club Auditorium	Location: Seminar Room 1	Location: Seminar Room 2
12:30pm	Chair: Ian Johnson , University of Sydney	Chair: Katherine Bode , Australian National University	Chair: TBC
12:30pm	Lunch		
-	Location: Ground Floor Terrace /Lower Colonnade		
1:20pm			
1:20pm	2.1: Language & Text II	2.2: Archives & Collections I	2.3: BOF: Digital Humanities and the Tyranny of Distance
-	Location: University Club Auditorium	Location: Seminar Room 1	Location: Seminar Room 2
2:50pm	Chair: Sydney Shep , Victoria University of Wellington	Chair: TBC	
2:50pm	Afternoon Tea		
-	Location: Ground Floor Terrace /Lower Colonnade		
3:20pm			
3:20pm	3.1: Language & Text III	3.2: Archives & Collections II	3.3: BOF: The Altruistic Crowd – The Ethics of Using Social Data in Humanist Research?
-	Location: University Club Auditorium	Location: Seminar Room 1	Location: Seminar Room 2
4:50pm	Chair: Nick Thieberger , University of Melbourne	Chair: Linda Barwick , University of Sydney	
5:00pm	Visualisation & Poster Reception (Sundowner)		
-	Location: HIVE & John Curtin Gallery		
7:15pm	Buses depart from the University Club for the John Curtin Galley, and return once the event is completed (approximately 7:30pm)		

Thursday, 20 March 2014 (Conference Day 2 - summary)

8:45am Registration
- **Location:** University Club Ground Floor Terrace
9:15am

9:15am **Keynote - Dr Sarah Kenderdine**
- Location: University Club Auditorium
10:30am Chair: Professor Hugh Craig, University of Newcastle

10:30am **Morning Tea**
- Location: **Ground Floor Terrace /Lower Colonnade**
11:00am

11:00am	4.1: Digital History & Virtual Heritage I	4.2: Network Analysis & Discovery	4.3: BOF: Building DH Community on the Regional Scale - Notes Towards a Multimodal Regional Digital Cultures Collaboratory
-	Location: University Club Auditorium	Location: Seminar Room 1	Location: Seminar Room 2
12:30pm	Chair: Jane Hunter , The University of Queensland	Chair: Sydney Shep , Victoria University of Wellington	Chair: TBC

12:30pm **Lunch**
- Location: **Ground Floor Terrace /Lower Colonnade**
1:20pm

1:20pm	5.1: Digital History & Virtual Heritage II	5.2: Infrastructure & Collaboration	5.3: BOF: All of the Records, All of the Time
-	Location: University Club Auditorium	Location: Seminar Room 1	Location: Seminar Room 2
2:50pm	Chair: Erik Champion , Curtin University	Chair: Paul Arthur , University of Western Sydney	

2:50pm **Afternoon Tea**
- Location: **Ground Floor Terrace /Lower Colonnade**
3:20pm

3:20pm	6.1: Digital History & Virtual Heritage III	6.2: Geo-Spatial Methods	6.3: BOF: Digital Humanities Afield
-	Location: University Club Auditorium	Location: Seminar Room 1	Location: Seminar Room 2
4:50pm	Chair: Simon Burrows , University of Western Sydney	Chair: Jo Hawkins, University of Western Australia	

6:30pm **Optional book reading & signing with author Alexander McCall Smith**
UWA Octagon Theatre, UWA Crawley Campus

Friday, 21 March 2014 (Conference Day 3 - summary)

8:45am	Registration		
-	Location: University Club Ground Floor Terrace		
9:00am			
9:00am	Keynote – Dr Anthony Beavers		
-	Location: University Club Auditorium		
10:15am	Chair: Professor Harold Short, University of Western Sydney		
10:15am	Morning Tea		
-	Location: Ground Floor Terrace /Lower Colonnade		
10:45am			
10:45am	7.1: Visualisation I	7.2: Music	7.3: BOF: Debate - That
-	Location: University Club	Location: Seminar Room 1	Literary Studies Needs
12:15pm	Auditorium	Chair: Harold Short,	More Graphs, Maps, and
	Chair: Jane Hunter, The	University of Western	Trees
	University of Queensland	Sydney	Location: Seminar Room 2
12:15pm	Lunch		
-	Location: Ground Floor Terrace /Lower Colonnade		
1:15pm			
1:15pm	8.1: Visualisation II	8.2: Literature & Literacies	8.3: BOF: Beyond
-	Location: University Club	Location: Seminar Room 1	Digitisation - Preparing
2:40pm	Auditorium	Chair: Jason Ensor,	Non-Digital Resources for
	Chair: TBC	University of Western	the Digital World and
		Sydney	Curation of Digital
			Resources in the GLAM
			Sector
			Location: Seminar Room 2
2:40pm	Afternoon Tea		
-	Location: Ground Floor Terrace /Lower Colonnade		
3:00pm			
3:00pm	9.1: Visualisation III	9.2: Digital Editions &	9.3: Digital Evidence
-	Location: University Club	Editing	Location: Seminar Room 2
3:50pm	Auditorium	Location: Seminar Room 1	Chair: TBC
	Chair: Deb Verhoeven,	Chair: Katherine Bode,	
	Deakin University	Australian National	
		University	
3:50pm	Closing Remarks		
-	Location: University Club Auditorium		
4:00pm			

Pre-conference workshop details

Monday 17th March 2014

9:00 – 17:00

Workshop – **Computation Based Discovery in Literary Language**

Hugh Craig, Jack Elliott & Brett D. Hirsch

This one-day course offers an introduction to the statistical analysis of literary texts, focusing on exploratory analysis rather than authorship attribution. We will spend most of the day learning and applying some techniques using software and text sets provided for use in the workshop and afterwards. No prior experience with statistics or with stylistics is necessary to participate.

The workshop outline is as follows:

- Basic introduction to computational stylistics
- Using simple word frequencies and conducting univariate analysis (e.g., correlations and t-tests)
- Applying more advanced methods of multivariate analysis (e.g., Principal Component Analysis, Zeta, Information Theory measures)
- Checking the validity of results
- Discussion

Workshop learning objectives:

- After completing the session, the attendee will be able to:
- Better understand a variety of statistical and computational methods and their application in the analysis of literary texts
- Apply methods of univariate and multivariate analysis to literary texts
- Critically assess the validity of results of such analyses

This workshop costs \$50.00 per person and prior [registration](#) is required.

Date: 17th March 2014

Who should attend this workshop:

This workshop will appeal to researchers and graduate students in literary studies and related disciplines.

Prerequisites:

Participants must provide their own laptop. Any operating system is acceptable, provided the most stable version of Java has been installed. To allow installation of statistical software, participants should also ensure that their laptop has a functional USB port.

14:00 – 17:00

Workshop – **Create beautiful scholarly maps with TileMill**

David Flanders & Steve Bennett

Do you have data that needs to be displayed on a map? This is a hands-on workshop teaching researchers how to create maps from research data with TileMill, an easy-to-use web-based cartography tool. It is useful for any researcher who needs to produce maps showing geographically-oriented data sets (locations of language communities, cultural artefacts, historical events) – or even just a thematically appropriate base map to incorporate into a website. Researchers will be given the tools to have full aesthetic control over the maps they create, for presentations, posters, papers, blog posts or web sites. Humanities researchers very often have geospatial datasets (although they don't call them that) they wish to communicate. One common approach is outsourcing the mapmaking process to a GIS specialist with a conservative result and little opportunity to play, innovate or iterate. The other common approach is to drop marker pins on a Google Map, resulting in a very impersonal, cluttered, ugly, thematically wrong map (freeways in 1860). TileMill lets you become the cartographer, with full control over how every element is rendered – or not.

This workshop is an intensive, condensed form of a course normally given over three evenings. Hence some materials will be given as an overview with resources for further self-learning.

The workshop outline is as follows:

- overview of TileMill capabilities (15 minutes)
- basic map styling using CartoCSS (60 minutes, including several brief exercises)
- loading data and using geocoding to convert addresses to locations (45 minutes)
- exporting maps for papers, posters (15 mins)
- creating webmaps for blog posts, web sites (15 mins)
- Q&A, additional material as requested throughout (30 mins)

Workshop learning objectives:

After completing the session, the attendee will be able to:

- create a map from appropriately prepared research data
- use CartoCSS, a cartography formatting language, to create an aesthetically pleasing map
- publish maps as either static images or dynamic web maps

This workshop costs \$150.00 and prior [registration](#) is required.

Date: 17th March 2014

This workshop will appeal to postgraduate students and academics. Anyone is welcome from industry, universities and government.

Prerequisites:

Participants must provide their own laptop. Those without laptops should form pairs or threes to participate in the exercises.

14:00 – 17:00

Workshop –Boosting your research profile through your research data

Adrian Burton

Research data is increasingly recognised by both institutions and researchers as a valuable asset, one well worth managing, citing and publishing. This interactive workshop will explore how researchers and institutions can optimise the impact and influence of research data products. Participants will use their laptops or ipads to determine the current extent of discoverability of their own data products, then generate an action list of ways to extend their research profile using current and future research data products. Humanities has rich and complex research data and research materials which are well suited to emerging publishing and metrics opportunities. Presenters for this workshop will include the Australian National Data Service as well as Humanities researchers and data managers. The workshop will first establish a definition of what constitutes humanities digital research data and then take a kaleidoscope approach to exploring elements of a range of new research data publishing paradigms to ensure that research data is as visible and cited as possible.

Elements will include:

- using data citation to contribute to your publication metrics and altmetrics exploiting emerging global trends such as the Open Researcher ID, DataCite and the Thomson Reuters Data Cit Index
- data management basics which encourage discovery and reuse (e.g. metadata, licensing, ethics, digital object identifiers)
- hot-housing citation and collaboration opportunities through linking data, research materials, journal publications, associated software, derived datasets, conference proceedings, grants etc
- connecting your various scholarly identifiers for greater visibility in global scholarly information systems

Humanities is a broad and global discipline and as such, this workshop will include both national and international examples and practices. Participants will examine their own research data and materials in the context of what is discussed and also contribute to the body of workshop knowledge through the perspective of their specific discipline.

The workshop outline is as follows:

1. Using data citation to contribute to your publication metrics and altmetrics (40mins)
2. Exploiting emerging global trends such as the Open Researcher ID, Datacite and the Thomson Reuters Data Citation Index (20mins)
3. Data management basics which encourage discovery and reuse (e.g. Metadata, licensing, ethics, digital object identifiers): hands-on component (40mins)
4. Hot-housing citation and collaboration opportunities through linking data, research materials, journal publications, associated software, derived datasets, conference proceedings, grants etc (20mins)
5. Connecting your various scholarly identifiers for greater visibility in global scholarly information systems: hands-on component (30mins)

Workshop learning objectives:

After completing the session, the attendee will be able to:

1. Recognise what their researcher profile is currently and can explain how various tools can be used to optimise their researcher profile
2. Apply multiple data management approaches to ensure their research data is managed, discoverable, able to be cited, and connected with their other research outputs
3. Appraise when various data management, researcher ID and publishing opportunities will be most appropriate to maximise the impact and reach of their research

This workshop is free of charge – however prior [registration](#) is required.

Date: 17th March 2014

Who should attend this workshop:

The following people would benefit from this workshop:

- Higher Research Degree students
- Early Career Researchers
- Researchers, Data Manager & Librarians

Prerequisites:

Participants are asked to identify a research dataset or collect they are familiar with, and bring access details or have extensive knowledge of what it contains. Participants are also requested to provide their own laptop computer.

Tuesday 18th March 2014

9:00 – 12:00

Workshop –#altac Track: Strategies to imagine and build alternative academic careers'
Jason Ensor & Jo Hawkins

#Altac careers have been described as off the tenure track, but within the academic orbit. These positions are serviced by “hybrid humanities scholars” who work in a diverse range of institutions including universities, cultural heritage bodies, libraries, museums, academic publishing, and the public sector. Roles often include a combination of administration, project management, teaching and research work.

The workshop outline is as follows:

This free, 3-hour workshop will explore strategies and tools to help PhD candidates and early career researchers imagine and build alternative academic careers. It will include a panel session with several practicing “alt-academics”, hands-on activities and plenty of lively discussion and debate.

We’ll be asking questions such as:

- How can PhD students better locate their research and work in an employable context?
- How can PhD students prepare for alternative academic careers during and after their candidature?
- What are the main opportunities and challenges associated with alternative academic careers?

This workshop is free of charge – however prior [registration](#) is required.

Date: 18th March 2014 (Morning)

Who should attend this workshop:

Postgraduate students and early career researchers.

Prerequisites:

No prior knowledge is required. Participants are encouraged to bring a laptop or a tablet.

9:00 – 12:00

Workshop –Federated Archaeological information Management Systems
Brian Ballsun-Stanton & Ian Johnson

Federated Archaeological information Management Systems (FAIMS) is a National eResearch Collaborative Tools and Resources (Australia) funded project to produce a comprehensive information system for archaeology and related fieldwork disciplines. Through community engagement it has developed flexible, robust and extensible tools for acquiring, refining, and archiving archaeological and related data. It allows data from field and laboratory work to be born digital using mobile devices, processed in web applications (local or online), and published online through a data repository. Means for facilitating the production of semantically, as well as syntactically, interoperable datasets have been built into the application at multiple points in the data lifecycle. Since the needs of archaeological fieldwork and research vary – and because many earlier efforts to construct archaeological data resources failed from being overly prescriptive – the project has developed the core of a federated, open-source system, encouraging the growth of an extensible range of options at each stage of data management. Solving the general information systems problem in archaeology has led to a system with potential applications in other disciplines, such as capturing historical data during oral interviews or archival research. The information management resources developed by this project include three broad categories of components, corresponding to the major stages in the life-cycle of archaeological data:

1. a new Android/Linux mobile platform for the creation of geospatial, multimedia, and structured data, including remote and offline deployments.
2. a web application (Heurist) for data refinement and analysis.
3. an online repository (based upon the Digital Archaeological Record) for storage and dissemination of data.

This workshop builds follows early demonstrations of the system at the Computer Applications in Archaeology conference (Perth, March 2013) but now presents these tools (especially the mobile platform) in a much more mature state.

The workshop outline is as follows:

Participants in this workshop will engage in hands-on demonstrations of new production releases of software developed by the FAIMS project, learning the system and providing feedback on various elements of design and performance. Participants will have the opportunity to use mobile devices loaded with the mobile application and access online tools for data refinement and archiving. Participants will learn how to customise and deploy the FAIMS mobile platform, which has been designed to solve the general problem of capturing a variety of data in a range of settings. The entire data lifecycle will be covered, from project implementation through data capture to refinement and archiving of data. Specifically the workshop will be presented as follows:

1. Introduction to the project and workshop orientation (10 min).
2. Introduction to Heurist for customising /deploying mobile apps data capture modules (10 mins).
3. Hands-on use of Heurist to create a data capture module for the FAIMS mobile application plus deployment of that module to mobile devices (50 mins).
4. Demonstration of the FAIMS mobile application (10 min).
5. Hands-on use of the FAIMS Android mobile application for geospatial, multimedia, and structured data capture using the module developed in (3) above (40 minutes).
6. Demonstration of Heurist for data refinement (10 mins).
7. Hands-on use of Heurist to import and modify mobile application data from (5) above (20 mins).
8. Demonstration of the FAIMS online repository (10 mins).
9. Hands-on use of the online repository, including upload, search, and retrieval (20 minutes).

Feedback received during this workshop will be used to inform future development of these tools and resources. We will have specific questions for workshop participants, but also encourage free responses and suggestions. Further details about the applications and services being demonstrated for feedback will be made available in early 2014.

Workshop learning objectives:

After completing the workshop, participants will be able to:

1. Use the FAIMS Android mobile application to capture geospatial, multimedia, and structured data.
2. Pass data from the mobile device to the data refinement web application (Heurist) and the repository (tDAR).
3. Manipulate data using Heurist.
4. Develop and customise simple data schemas and user interfaces in Heurist for deployment to the mobile platform for data capture.
5. Assess the applicability of FAIMS software to their information management needs.

This workshop is free of charge – however prior [registration](#) is required.

Date: 18th March 2014 (Morning)

Who should attend this workshop:

Academic and professional archaeologists at all levels; other researchers who need to capture data in the field and refine it afterwards (e.g., historians conducting oral interviews or undertaking archival work). No prior knowledge or skills are required; the purpose of this workshop is to solicit feedback from as wide a range of archaeologists as possible – we look forward to hearing from users with previous experience of mobile device applications or archaeological databases and encourage their participation, but such experience is not a prerequisite.

Prerequisites:

No particular prior knowledge required. Participants are encouraged to bring their own laptops and Android 4.0+ mobile devices (some will be provided).

9:00 – 12:00

Workshop –Automated 3D Model Reconstruction from Photographs

Paul Bourke

Photogrammetry is the traditional name given to the derivation of some 3D quantity derived solely on the basis of a collection of photographs. In recent years this has been an intensive area of research and the quality of the algorithms is reflected in this effort. The state of the art is currently that one can create reasonable quality 3D models from a collection of photographs without a significant amount of domain knowledge, specialist hardware and software, or with the need to place markers in the scene or follow rigorous calibration procedures.

This workshop will discuss the theory, introduce the current software solutions/pipelines, make reference to camera and techniques that result in an optimal chance of successful reconstructions, and present some of the post processing requirements and tools. In short, the workshop will aim to provide a complete introduction to the subject. The emphasis will be towards 3D capture of heritage objects and this will be the topic of most of the examples.

The workshop outline is as follows:

- Introduction, outcomes and motivation
- History and alternative technologies
- Main software aspects and titles, pipelines
- Photography: lenses, shooting styles
- Case Study 1: 2.5D reconstruction example
- Geometry processing, working with meshes, file formats
- Case study 2: Full 3D example
- Other related topics: resolution, relighting, analysis
- Limitations: occluders, scene movement, shadows
- Case study 3
- Additional applications/topics including rapid prototyping
- Further reading, references, and discussion

There will optionally be an opportunity for attendees to experiment by taking photographs and testing reconstructions using software provided.

Workshop learning objectives:

At the end of the workshop attendees should feel confident in applying and experimenting with this exciting new technology. By attending the workshop the common pitfalls can be avoided and attendees will be aware of the key concepts, familiarity with relevant software tools, realistic expectations and limitations, knowledge of applications and research where 3D reconstruction is being employed.

This workshop is free of charge – however prior [registration](#) is required.

Date: 18th March 2014 (Morning)

Who should attend this workshop:

This workshop is targeted at those who wonder if they could make use of 3D model reconstruction from photographs and would like a thorough introduction to the subject. The workshop is ideally suited to those who are aware that photogrammetry is being used to reconstruct objects but is not aware of the state of the technology, what is possible and what is not.

Prerequisites:

No particular prior knowledge required. If participants wish to experiment they can bring their own cameras but they need to discuss this with the presenter beforehand.

9:00 – 12:00

Workshop –A taste of HuNI

Toby Burrows & Alex Hawker

This workshop is designed to introduce humanities researchers to the HuNI Virtual Laboratory. Researchers will be given an introduction to the contents and capabilities of the HuNI VL, and its relationship to the various contributing datasets. They will learn how to create their own account in

HuNI and use it to build and share collections of data relevant to their research. They will also learn how to annotate entities in HuNI to show relationships between them, and how to export information from HuNI.

After completing the workshop, attendees will be able to start using the HuNI Virtual Laboratory as an integral part of their research. The workshop is not intended to be an in-depth look at the technical architecture and functionality of the Virtual Laboratory, and is not designed for technical experts.

The workshop outline is as follows:

- Introduction: overview of the HuNI Virtual Laboratory, its architecture, data and capabilities. (30 minutes)
- Searching and browsing the HuNI VL, navigating the contents of the VL, interpreting and using search results. (20 minutes)
- Creating and managing your user account. (10 minutes)
- Creating, managing and sharing your own collections within the VL. (20 minutes)
- Building relationships between entities in the HuNI data aggregate. (20 minutes)
- Exporting information from the HuNI VL. (20 minutes)

With the exception of the introduction, the workshop will be hands-on activities for all participants. This workshop is free of charge – however prior registration is required.

Workshop learning objectives:

After completing the session, the student will be able to:

1. Understand the scope and capabilities of the HuNI VL
2. Discover resources in the HuNI VL
3. Set up and manage their personal HuNI account
4. Build and share personal collections of HuNI data
5. Work with HuNI entities to create relationship links
6. Export information from HuNI

This workshop is free of charge – however prior registration is required.

Please use the following [registration if you wish to attend the morning session](#).

Please use the following [registration if you wish to attend the afternoon session](#).

Date: 18th March 2014 (Morning and repeated in the Afternoon)

Who should attend this workshop:

Humanities researchers, postgraduate students, undergraduate students, and independent researchers.

Prerequisites:

Students are expected to have some prior knowledge and familiarity with searching and using humanities databases. Participants can use their own laptops/devices with wireless access during the workshop.

14:00 – 17:00

Masterclass – Computational philosophy and the moral implications of automated decision making
Anthony Beavers

The Masterclass is an opportunity for postgraduate students and early career researchers to meet and discuss specific research topics with a distinguished scholar. Participants will be invited to participate within the framework of the topic as part of a discussion.

Masterclass topic: Computational philosophy and the moral implications of automated decision making

Abstract: In the emerging area of cognitive robotics and other cognitive computer applications machines make actual ethical decisions themselves, thus taking humans out of the loop, or minimally helping humans make value decisions, which raises some challenging questions about what humans are doing when they make value decisions.

Host: The Masterclass is being hosted by Dr. Anthony F. Beavers, Professor of Philosophy and Director of the Digital Humanities Lab, The University of Evansville. Tony Beavers works in the developing area of computational philosophy, an approach to the discipline that involves using computers to make philosophical discoveries that are not readily available with traditional argumentative methods and that also tests philosophical theories for computational tractability. Tony is also a keynote speaker at DHA 2014. Tony's full bio is available from [here](#).

Registration: The Masterclass is free, and registration is essential.

To register, select the Masterclass please visit the [registration site](#).

Location: iVEC@ECU Motion Capture Facility, Building 13, Mount Lawley Campus.

Time: 2:00pm – 5:00pm

Social event: The Masterclass will be followed by a musical evening reception (from 5:00pm onwards) at ECU.

14:00 – 17:00

Workshop – **A taste of HuNI** (Repeat of morning workshop, see above)
Toby Burrows & Alex Hawker

Main Conference - Wednesday, 19th March 2014

- 8:30 – 9:00 Registration
- 9:00 – 9:15 Opening by Professor Paul Johnson, Vice Chancellor of The University of Western Australia followed by traditional Welcome to Country by Barry McGuire
- 9:15 – 10:30 **Keynote** - Professor Neil Fraistat
Maryland Institute for Technology in the Humanities, University of Maryland, USA

Neil Fraistat is Professor of English and Director, Maryland Institute for Technology in the Humanities (MITH), University of Maryland, USA. Neil currently chairs the international Alliance of Digital Humanities Organizations (AHDO), is Co-Founder and Co-Chair of centerNet, an international network of digital humanities centers, and is Vice President of the Keats-Shelley Association of America. Fraistat is Co-Founder and General Editor of the Romantic Circles Website and has published widely on the subjects of Romanticism, Textual Studies, and Digital Humanities in various articles and in the ten books he has authored or edited.

Neil's most recent publications are *The Cambridge Companion to Textual Scholarship* and Volume III of *The Complete Poetry of Percy Bysshe Shelley*. He has been awarded both the Society for Textual Scholarship's biennial Fredson Bowers Memorial Prize and the biennial Richard J. Finneran Prize, the Keats-Shelley Association Prize, honorable mention for the Modern Language Association's biennial Distinguished Scholarly Edition Prize, and the Keats-Shelley Association's Distinguished Scholar Award.

Location: University Club Auditorium
Chair: Professor Paul Arthur, University of Western Sydney / President of aaDH

11:00- 12.30 **Sessions**

1.1: Language & Text I

Location: Seminar Room 1

Improving Access to Recorded Language Data

Simon Musgrave^{1,3}, Linda Barwick^{2,4}, Michael Walsh^{2,5}

¹Monash University, Australia; ²University of Sydney; ³Australian National Corpus;
⁴PARADISEC; ⁵AIATSIS

Researchers in different disciplines collect and store data which includes human language recorded in real time. Discovery of such data should be easy across disciplines, but is currently impeded by different disciplinary approaches and standards. For example, a linguist may have collected recordings of songs performed by speakers of the language they are studying; these recordings are stored in an archive intended primarily for other linguists, and a musicologist may not easily discover the resource even though it might be very relevant to their research. This paper will discuss the work of a recently formed Working Group within the Research Data Alliance which aims to address this problem by working towards standardisation of metadata

elements in two areas: codes for identification of languages and language varieties, and categories for describing the content of resources.

For language identification, ISO639-3 provides a set of three letter codes to identify languages. But this is not unproblematic for a variety of reasons. Firstly, it is not adopted everywhere; for example the digital collections of the Australian Institute of Aboriginal and Torres Strait Islander Studies use a different set of identifiers and this example also shows two other problems for language identification. The divisions recognised by ISO639-3 do not always align with expert understanding. This has been a particular issue for Australian languages, with a number of change requests filed with the registration authority for ISO639-3. A number of these changes relate to delineating languages from linguistic entities below that level (such as dialects) and above that level (such as macrolanguages and language families). Proposals for identification of entities at different levels of granularity are being considered within the ISO process; the Working Group aims to ensure that expert input to these processes is maximised, that the principles underlying the ISO639 standard sets have a sound linguistic basis, and that registration and revision processes are consistent and transparent. Our assumption is that progress with these issues will lead to more consistent use of the standard by archives and repositories. Existing metadata schemas (e.g. IMDI, OLAC) include a vocabulary for describing the genres represented in linguistic resources, but these do not necessarily correspond to needs of different disciplines. Consultation across different research communities is needed to establish the range of resource types which need to be covered and vocabularies for describing that range. The Working Group will implement the results of this consultation by creating a set of metadata elements within the frameworks of the Component Metadata Initiative (CMDI) and the ISOCat data category registry. CMDI allows for the use of common metadata elements across different sites without imposing a rigid metadata scheme, while the ISOCat framework ensures that the semantics of (meta) data elements are explicit and accessible.

We hope that the activities of the Working Group will lead to improved discovery and access for researchers across disciplines who work with recorded language data as well as improved possibilities for inter-repository data exchange.

Research outcomes of encoding vocabularies of Australian Indigenous languages

Nick Thieberger, Conal Tuohy
University of Melbourne, Australia

Collections of Australian language vocabularies that have been created as digital files have not, to date, been able to be linked together or to be explored as a large dataset. Individual comparative linguistic projects have established stand alone lexical databases and Native Title cases have often required the preparation of historical vocabularies for comparison over time. But there has been no effort to collect these various sources into a common format and there is currently no repository for digital files related to Australian Indigenous languages. This paper discusses the Digital Daisy Bates project which is using the Text Encoding Initiative's schema (TEI) to encode text from archival sources, and to create new research outcomes by exploring the encoded material with novel techniques. Some 23,000 pages of these manuscripts are stored at the National Library of Australia and the Barr Smith Library in Adelaide and 3,300 pages have so far been keyboarded in TEI format and geocoded. This project will allow the navigation of the diverse vocabularies on a map, and will show dialect

variation within large family groups in Western Australia. Ontologies for people and placenames will provide further points of entry to the data, and will allow linking to external authority files, including the NLA's People Australia. The structured data will also be converted to RDF to build a linked data set. As the underlying material is in a re-usable format it will be possible to publish electronic editions of the manuscripts, output sets of wordlists for use in schools, and to do comparative analysis of the vocabularies to infer locations (as in Nash 2002, and more recently Embleton et al 2013).

This paper will be of interest to linguists and to other researchers working with structured data in the form of archival records.

Sheila Embleton, Dorin Uritescu and Eric S. Wheeler. 2013. Defining dialect regions with interpretations: Advancing the multidimensional scaling approach *Lit Linguist Computing* 28: 1 13-22.

Nash, David. 2002. Historical linguistic geography of south-east Western Australia, pp.205-30 in *Language in Native Title*, ed. by John Henderson & David Nash. Canberra: AIATSIS Native Title Research Unit, Aboriginal Studies Press.

Profiling developmental patterns of learner language

Mariko Abe¹, Yuichiro Kobayashi²

¹Chuo University, Japan; ²Japan Society for the Promotion of Science

Learner corpora, digitized learner performance data, have the potential to clarify our understanding of interlanguage development. However, few attempts have been made to provide persuasive results about language development that can be applied in language teaching and assessment materials (Pendar & Chapelle, 2008). Despite early work in Second Language Acquisition (SLA), relatively few researchers have been concerned with the development of oral proficiency and have profiled the interlanguage development from multiple perspectives. More specifically, the number of targeted linguistic features in previous SLA studies is limited (Biber, Conrad, & Reppen, 1998). Consequently, it is still not clear which linguistic features identify different developmental stages, and this results in an insufficient understanding of comprehensive overview of the characteristics of interlanguages.

This study aims to profile developmental patterns of second language (L2) spoken English across eight oral proficiency levels of Japanese EFL learners. In order to achieve this purpose a methodological approach combining learner corpus, language processing techniques, and multivariate statistical analysis was used to identify the frequent linguistic co-occurrence patterns from over one million running words of L2 spoken English. The frequencies of 58 linguistic features on vocabulary, parts-of-speech, semantic categories for the major word classes, and grammatical characteristics, which have been shown to distinguish text and style variation in Biber (1988), were used for the analysis.

As the frequencies of specific linguistic features have a central role in this study, the largest spoken learner corpus in Japan, the National Institute of Information and Communications Technology Japanese Learner English (NICT JLE) Corpus (Izumi, Uchimoto, & Isahara, 2004), which consists of spoken interview transcripts of over 1,200 learners who have taken the Standard Speaking Test (SST), was chosen. This oral English test was constructed on the basis of the American Council on the Teaching of Foreign Language Oral Proficiency Interview (ACTFL OPI), and it was designed to evaluate Japanese EFL learners. The individual scripts are coded with eight oral proficiency levels by certified raters.

After calculating the frequencies of 58 linguistic features, we employed a multivariate statistical analysis, heat map with hierarchical clustering. This statistical analysis is powerful in dealing with a large number of variables and in visualizing the complex interrelationships

among oral proficiency levels, interrelationships among linguistic features, and the association patterns between levels and features.

As a result, the eight oral proficiency groups were distinguished in four clusters that correspond with the levels of ACTFL OPI: Novice (SST level 2 and 3), Intermediate Low (SST level 4 and 5), Intermediate Mid & High (SST level 6, 7, and 8), and Superior (SST level 9). Lower-level learners tend to use personal pronouns, nouns and predicative adjectives, while higher-level learners use THAT relatives, WH-relatives, and emphatics. We suggest the possibility of these linguistic features being used as indices of L2 spoken development. The results of the present study can contribute to complement an insufficient profiling of the L2 spoken language use.

Biber, D. (1988). *Variation across speech and writing*. New York: Cambridge University Press.

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

Izumi, E., Uchimoto, K., & Isahara, H. (2004). *A speaking corpus of 1200 Japanese learners of English*. Tokyo: ALC Press.

Pendar, N., & Chapelle, C. (2008). Investigating the promise of learner corpora: Methodological issues. *CALICO Journal*, 25(2), 189-206.

Identifying discipline-specific expressions based on institutional repository

Yuichiro Kobayashi¹, Shosaku Tanaka², Yoichi Tomiura³, Yoshinori Miyazaki⁴, Michio Tokumi³
1Japan Society for the Promotion of Science, Japan; 2Ritsumeikan University, Japan; 3Kyushu University, Japan; 4Shizuoka University, Japan

From the early 1960s, English for the Specific Purposes (ESP) has become one of the prominent fields of English teaching (e.g. Hutchinson and Waters, 1987). It aims to provide learners with academic English skills in a context of tertiary education. In terms of English vocabulary, the Academic Word List (AWL) contains 570 word families selected from academic texts (Coxhead, 2000). The list has been widely known to both academic English teachers and researchers, but does not cover context-bound or topic-dependent vocabulary frequently used in each discipline and sub-discipline.

The most significant words in each academic context differ depending on the discipline, university, laboratory or researcher. Therefore, it is necessary to compile a large corpus which covers a wide range of academic domains to create a more appropriate word list for researchers in different academic fields. What is more, it is important for language learners to acquire discipline-specific expressions to achieve native-like performance in an ESP context.

The purpose of this study is to identify discipline-specific expressions, especially multi-word expressions, using institutional repositories and natural language processing techniques. An institutional repository is an online collection of the intellectual output of a research institution. The recent trend of constructing institutional repository in Japan allows researchers to share the same academic resources. The present study demonstrates the effectiveness of creating a multi-word list specific to an institution.

We used a gapped n-gram approach for the identification of discipline-specific expressions. An 'n-gram' is a contiguous sequence of items, such as words or part-of speech tags, and a 'gapped n-gram,' or 'skipgram,' is a refinement of the n-gram approach designed to detect non-contiguous item associations (e.g. Cheng, Greaves, and Warren, 2006). Our algorithm for extracting gapped n-grams is based on the method proposed by Kozawa, Sakai, Sugiki, and Matsubara (2010). First, we filtered the texts in the institutional repository by the number of words to exclude too short or too long writings. Second, we used TreeTagger, an automatic

chunking program, to identify the basic phrase and clause structures. Third, we detected general patterns of multi-word expressions by lemmatizing words except for participial verbs and by replacing determiners and cardinal numbers into part-of-speech tags. Finally, we counted the number of gapped n-grams occurred in the writings collected from each academic disciplines. This method enabled us to effectively detect frequent patterns characteristic of each discipline. We applied the method to the QIR, the institutional repository of Kyushu University, using the repository of the Faculty of Information Science and Electrical Engineering (ISEE), and identified discipline specific expressions for each faculty at the university as an example of the automatic detection of gapped n-grams. We collected 229 writings in the QIR, and successfully obtained 1,061 frequent phrasal expressions specific to ISEE. We expect this study contributes to teaching English for the Specific Purposes and the application of digital library to linguistic research.

Cheng, W., Greaves, C., and Warren, M. (2006). From N-gram to Skipgram to Congram. *International Journal of Corpus Linguistics*, 11(4), 411-433.

Coxhead, A. (2000). A New Academic Word List. *TESOL Quarterly*, 34(2), 213-238.

Hutchinson, T., and Waters, A. (1987). *English for the specific purposes*. Cambridge: Cambridge University Pres.

Kozawa, S., Sakai, Y., Sugiki, K. and Matsubara, S. (2010). Automatic Extraction of Phrasal Expressions for Supporting English Academic Writing.

Proceeding of the 2nd International Symposium on Intelligent Decision Technologies (KES-IDT-2010), 485-493.

1.2: Print Cultures

Location: University Club Auditorium

Presentations

Mapping Print, Connecting Cultures

Simon Burrows, Jason Donald Ensor

University of Western Sydney, Australia

Twin advances in historical GIS and mapping and new bibliometric research techniques for recording, annotating and representing spatial-temporal data related to the production, dissemination and reception of books, have created new opportunities for understanding the dissemination of print and ideas within and across cultures. This paper will unpack the affordances and challenges of mapping print cultures across time and space; and the potential of this approach to help us perceive and understand in new and nuanced ways the transformative powers of print communication and the complex relations of audiences and texts across large expanses of time and space. The paper will review the state of attempts at historical GIS/mapping and the various digital approaches to bibliometric and reader reception before outlining a vision for how this work might be further developed, both on a global level and in specific historical contexts, including Australia and New Zealand. In doing so it will draw on the speakers' wide experience of digital mapping and historical work on the history of publishing, with special reference to Jason Ensor's involvement in the ANU 'Deepening Histories of Place: Exploring Indigenous Landscapes of National and International Significance' project and the long history of Australian publishing and novels, and the pioneering work of Simon Burrows and his collaborators on the highly acclaimed 'French Book Trade in Enlightenment Europe' project.

Mapping Printers' Lives and Letters

Sydney Shep, Flora Feltham, Sara Bryan

Victoria University of Wellington, New Zealand

The so-called 'spatial turn' in history has led to new conceptualisations of space, place, locus and site. Geospatial technology "articulates the idea of absolute Euclidean spaces quite well, but the socially-produced and continuously changing notion of place has to date proved elusive to digital representation except, perhaps, through photography and film" (Fisher and Unwin, 2005). If, as Doreen Massey has remarked, "places are constructed at a particular constellation of social relations meeting and veering at the particular locus" (1994, 28), then a spatial analysis of the book trade is essential to an understanding of the complex intercrossings between the production, circulation, and reception of the printed word. Using the example of New Zealander Robert Coupland Harding (1849- 1916) and his landmark journal, *Typo*, this interactive paper discusses how new geotemporal visualisation tools such as Neatline, HistoryPin, and BatchGeo amongst others help to illuminate the individual and networked biographies and bibliographies of nineteenth-century printers. By plotting Harding's extensive local and international connections, tracing the dissemination of his design principles around NZ, and charting his downward spiral from relative prosperity in Napier to bankruptcy in Wellington, we also engage with Tim Hitchcock's recent call to interrogate "how the evolution of the forms of delivery and analysis of text inherent in the creation of the online, problematizes and historicises the notion of the book as an object, and as a technology; and in the process problematizes the discipline of history itself as we practise it in the digital present" (<http://digitalhumanitiesnow.org/2012/01/academic-history-writing-and-itsdisconnects-by-tim-hitchcock/>).

Data Models and Services: Australian Literature and Print History

Rowan Brownlee², Kerry Kilner³, Ingrid Mason¹

¹Intersect Australia, Australia; ²University of Sydney; ³University of Queensland

Two domains of research are the focus for this discussion: Australian literature and print history and data modelling and information management. Research in Australian literature and print history requires access to relevant metadata and the works (physical and digital) to which they refer as well as the tools to analyse literary history. Access to open public access catalogues and databases, e.g. Trove, AustLit, and other selected and curated collections of digitised works have enhanced traditional scholarly research practice. There is further potential for improving research outcomes through the wider take up of digital scholarship and digital humanities methodologies. However, service and data models may need to change in order to appropriately support that work as discovery and linking across datasets becomes more crucial.

Three questions are posed:

- * What data do researchers need access to and why?
- * Who produces the data and how do researchers want to use it?
- * How can the data be repackaged, linked and re-presented to support enhanced use?

The authors of this paper examine methods of data modeling, translation and services that advance digital scholarship and digital humanities research from different perspectives: the digital humanities researcher (Kerry Kilner, AustLit), the research analyst (Ingrid Mason) and the digital project analyst (Rowan Brownlee).

Re-binding the book: material frames for unbound books

Tully Barnett

Flinders University, Australia

The rise of the eBook is frequently dated back to the work of Michael Hart in making a digital copy of the US Declaration of Independence available on the University of Illinois computer network in 1971. This became the first text in the Project Gutenberg text archive, which went online in 1994, and grew to 42 000 out of copyright texts over the next decades. The works were unbound: massive slabs of text rendered immaterial by the process. They could be copied and pasted, emailed, inserted into a word processing document or a html document or given other new frames or bindings – including new physical bindings by being reprinted. Archive.org's text followed suit, providing .txt files with ultimate flexibility but divorcing the work entirely from the material container of the book. This is consistent with the rhetoric of the rise of the immaterial over the material prevalent in the early days of the internet.

Recently, however, a plethora of apps and reading platforms have arisen that emphasise nostalgia for the printed text in the digital reading experience. Apps frame books with pages to be turned by finger swiping or clicking, and visualise the contents of the mobile library on a wooden shelf. Kindle has introduced page numbers indicative of the printed text rather than solely providing a percentage of reading completed.

This shift back to a material and nostalgic reading experience can be attributed to many factors; this paper considers three of them. Firstly, the Google Books project created digital texts through scanning and OCR but ultimately the digital archive version retained the image of the page rather than allowing the OCR text to supersede it. Where Project Gutenberg and Internet Archive provided only the electronic text of the work – whether it was created by OCR or by hand transcription – Google Books, and now the HathiTrust, provide the digital image of the scan, plus access to the OCR text for searchability, in the archive. This has altered the relationship between the reader and the text. Secondly, the rise of mobile reading has created a focus on the material framing of the book seen in the design of online and mobile book reading applications and platforms as readers respond to the feedback provided by the bound book experience. Thirdly, the validity of the concept of the digital experience as immaterial, as transcending embodiment in material substrates, has been debunked in multiple ways, and this plays out in the experience of reading digitally. This paper considers these factors in light of the affect and affordances of reading in the digital environment and argues that design principles have sought to re-bind the book for largely nostalgic purposes.

1.3: Online Engagement & Communities

Location: Seminar Room 2

Presentations

Web series, disability humour and the disability rights agenda

Katie May Ellis

Curtin University, Australia

User generated content – blogs, YouTube, Facebook and Twitter – provide valuable real time opportunity for people with disability to engage in media industries and shape the media

agenda. Along with mitigating various types of impairment through alternative formats, these platforms are inclusive of the disability rights agenda. The new forms of televisual culture – centring on YouTube and other video-sharing and distribution platforms – allow ‘ordinary’ people with disabilities to present videos and programs on their lives, as well as presenting a new kind of medium where non-traditional media actors (disability or human rights groups, for instance) can quite easily devise and circulate media content. Similarly, social media – through the vogue applications of Facebook and Twitter – have not only been at the heart of reshaping of media, they also allow disability to figure in media in distinctive and novel ways.

For example, web series are emerging as an important avenue for disability representation as well as employment. This paper looks at the appearance of web series as a result of converging digital technologies and the ways these new modes of communication and representation include and exclude people with disabilities. Preliminary research on disability, media and popular culture groups on Facebook identified the web series *My Gimpy Life* as an important program for people with disabilities seeking out new forms of media such as web series. This paper examines the techniques used in a number of web series that attempt to offer a different representation of disability than what is commonly seen on broadcast television to report on ongoing research into disability, digital television and online media.

This paper reports on the findings of a survey into people with disabilities’ prosumption (cf Bruns) of web content. I consider web series that offer representations of disability/ disability themes/ actors with disabilities such as *My Gimpy Life*, *The Retarded Policeman*, and *The Guild*. I explore the use of disability as it intersects with genre conventions within these texts to make a comment on the social inclusion of people with disabilities. For example, *My Gimpy Life* which follows Teal Sherer, an aspiring actress with a disability as she navigates the inaccessible world of Hollywood auditions, productions and night life confronts disabling social attitudes through disability humour (Haller 2010). The paper concludes with reflections offered by people with disabilities regarding the potential web series offers to create a more empowering and less stigmatising representation of disability and how they are using digital and networked media to engage in media production themselves.

Engaging the "Quaint" Creative Writing Classroom in the Social Media Marketplace

Susan Taylor Suchy
UWA, Australia

In spite of calls for more digital engagement and the fact that students are arriving on campus with digitally connected skills, most university-level creative writing classrooms are generally “low tech and quaintly humanistic”(1). We don’t appear to be incorporating the socially networked student experiences in the “quaint” creative writing classroom. One of the barriers to more engagement may be one of the hard-won “markers of professional difference,”(2) that is, the things that distinguish creative writing from other classes. This particular marker is that the creative writing classroom is not market-driven. By examining this issue (but not eliminating the marker), we can determine how to open the class up to more engagement, particularly in the area of social media and the social media marketplace. This paper examines the reasons why certain classrooms don’t engage, the terms “social media” and “social media marketplace” (in order to consider changes to the marketplace and possible benefits from a certain aspect of digital engagement), and offers some examples of engagement that honor

traditions and benefit the classroom by enhancing educational experiences without excessive cost or training.

(1) McGurl, M 2009 *The Program Era: Postwar Fiction and the Rise of Creative Writing*, Harvard University Press: Cambridge, MA (21)

(2) Ritter, K 20013 "Professional Writers/Writing Professionals: Revamping Teacher Training in Creative Writing Ph.D. Programs", *College English* 64(2), (208)

Digital Life Narratives, Trauma and the Ethics of Reception.

Kate Douglas

Flinders University, Australia

Recently Dr Anna Poletti (Monash) and I have been engaged in researching life narratives authored by young writers and artists. In our project we've found three distinct spaces in which young cultural producers construct and circulate autobiographical narratives: private written forms, public literary forms, and multimedia texts, which blur the boundaries between private and public. The interstices between so-called private and public narratives are of great interest to us. Of course, even when an author chooses to circulate his or her life narrative publically they cannot foresee the text's transit: who will read it and how it will be read. For example, a young microblogger writing trauma may not anticipate the possible negative implications of his or her once private story entering the public domain, where reception may include suspicion, public interrogation, even ridicule.

In this paper I consider a selection of social media and crowdsourcing projects and texts which target young writers and artists. I will then explore what an ethical reading and ethical scholarship of these youth-authored digital trauma texts might require. This case study reveals the need for deep, complex questioning and reflection when it comes to the study of apparently public, youth-authored life narrative. I am also interested in the ethical dilemmas which may arise through the deployment of trauma narratives within digital humanities scholarship.

Young People's Safety and Wellbeing in Digital Culture: Affordances and Infrastructures

Teresa Swirski¹, Philippa Collin²

¹University of Western Sydney, Australia; ²University of Western Sydney, Australia

This presentation describes a Young and Well Cooperative Research Centre project exploring key issues that affect young people's safety and wellbeing, plus the potential of social marketing campaigns to influence their online and offline behaviour. The affordances of persistence, replicability, scalability and searchability characterising 'networked publics' (boyd, 2011) highlight the novel challenges and opportunities young people now face. Illustrated is the participatory design approach utilised to develop a series of social marketing campaigns, as well as the longitudinal tracking and innovative digital data collection to measure reach, outcome and impact. Such a methodology helps to examine the complex social, material and temporal aspects of the Online Social Marketing for Safety and Wellbeing project: the interrelationship between youth participants, partner organisations, digital technologies, and wellbeing. The two campaigns developed so far are discussed in relation to Deuze's (2009) principal components of digital culture: participation, remediation and bricolage. The first campaign, Keep it Tame, is an interactive online video encouraging young people to reflect upon the repercussions of their peer communications; plus options for them to engage further with safety and wellbeing issues via a range of resources and service providers. The second

campaign, @ppreciate, is an online tool developed to foster positive communication among young people; messages can be selected, modified and shared across a range of social media platforms. The development and impact of these campaigns provide rich insights into the diverse ways in which young people imagine, improvise and innovate their practices amid digital culture. Such complexities are captured in Lievrouw and Livingstone's (2009) notion of 'new media infrastructures'; these are composed of artefacts, or devices, activities and practices, as well as social arrangements or organisational forms. Affordances and infrastructures help to frame the ways in which participation and knowledge translation unfold in this project between young people, sector and research partners, government, as well as digital media and online safety experts. Understanding more about these dynamics offers new insights into how we collectively imagine and shape digital cultural practices.

Boyd, d. (2011). Social Network Sites as Networked Publics. In Z. Papacharassi (Ed.) A Networked Self, Chapter 2, pp. 39-58. New York: Routledge.

Deuze, M. (2006). Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture. In: The Information Society 22(2), pp.63-75.

Lievrouw, L. A. & Livingstone, S. (2006) Introduction. In L. A. Lievrouw & S. Livingstone (Eds.). Handbook of new media: social shaping and social consequences – fully revised student edition, pp. 1-14. Sage: London.

13:30- 15:00 **Sessions**

2.1: Language & Text II

Location: Seminar Room 1

Presentations

Whole Genre Analysis with Weighted Gene Co-Expression Network Analysis

Jack Elliott

University of Newcastle, Australia

Romance accounts for 14.1 percent of all novels sold in the United States, and is the most popular form of genre literature. Selling into 111 different markets Harlequin is the largest publisher of romance world-wide, and its most successful category line, Harlequin Presents, offers the reader "pure romantic fantasy" - romance in an almost archetypal sense. Harlequin Presents is so popular that eight novels are published every month - a deluge of novels that forces researchers to rely on statistical analysis. The central question when studying any kind of genre fiction is the amount of variation in the novels. Does the category overwhelm the voice of the writer, or can writers use the category romance as a form of expression? Does the category romance vary over time, or is it stagnant? Finally, to what extent are the order of plot elements fixed and to what extent are they given free play within a novel? Adapting a technique - Zhang and Horvath's Weighted Gene Co-Expression Network Analysis - from the life sciences and applying it to all 1,400 digitally available Harlequin Presents novels, it can be shown that authorship is the dominant feature of the category, that the category changes in response to financial shocks at Harlequin, and that most elements such as setting and declarations of love have a fixed location in the category romance.

Unsupervised clustering of Weighted Gene Co-Expression Network data arranges each novel in the imprint by author, demonstrating authorship provides the organizing principle for the category. This is surprising, given that critics such as Leslie Rabine believe that the category romance overwhelms authorial voice. A further surprise is that authorship even overpowers

distinctions such as the sub-category (Sheikh novel, marriage-of-convenience plot and so on) or miniseries (collections of novels written by different authors incorporating the same characters). Intriguingly, this judgement is made based not purely on function words, but also on words to do with the type of hero (writers specialize in cynical, royal or Mediterranean heroes), inquit tags (many writers employ inquit tags more distinctive than "he said" or "she said") and use of adjectives.

Although there is a perception that category romances do not change, clustering the data by year produces two distinctive periods - an early period from 2000 to 2004, and a late period from 2005 to 2013. This change is a result of a financial shock at Harlequin in 2004, which forced Harlequin to revamp their categories. This shift demonstrates the publisher is in overall control of the direction of the imprint. Authors are, however, critical to the process of change, as Harlequin implement their vision through re-tasking and redeploying their authors.

Correlating Weighted Gene Co-Expression Analysis data to novel segment reveals that three key elements have a fixed order within category romance novels. Although these three elements - the setting, the sex-scene and the declaration of love - are restricted in where they appear, authors typically employ different strategies to realise these narrative elements.

Harlequin Presents, the most popular type of category romance, is regular in the ordering of its plot elements. The exact content of these elements varies dramatically by author - so much so that authorship serves as an organising principal for the category. The publisher, Harlequin, are in control of the overall direction of the category, a control they manage through their authors. This study demonstrates that even unsupervised statistical analysis using large data sets can describe the interrelation of authorship, genre and market forces more precisely than anyone had imagined.

Computational Stylistics and Cognitive Grammar: linking computational and contextual approaches in a study of early modern drama

Louisa Jane Connors

University of Newcastle, Australia

Burrows style computational stylistics has been successfully applied to a range of questions concerning textual features such as genre, authorship, and historical period. It is less common for the techniques associated with the Burrows method to be applied to the consideration of interpretive questions. This paper situates the methods and techniques of computational stylistics within a theoretical framework that is ground in Cognitive Grammar. It argues that when insights from Cognitive Grammar are brought to bear on a stylistic study, the results provide insights that complement traditional interpretive analyses and counter criticisms of the method arising from more orthodox approaches. This paper reports on the results of a study of sixty early modern tragedies printed between 1580 and 1640. The study compares twelve closet tragedies loosely associated with Mary Sidney, Countess of Pembroke, with forty-two tragedies written for the commercial stage, before exploring one of the closet texts – Elizabeth Cary's *The Tragedy of Mariam* – in more detail.

Traditional critical approaches to the study of Elizabeth Cary's *Mariam* have produced contradictory readings of the text and do not explain the many paradoxes that emerge through close analysis. It is difficult, for example, to align arguments that the text rejects the notion of performance with the ways it gestures towards the stage and dramatic presentation. Dramatic elements in the play also appear to contradict the common assumption that the

intellectual and literary associations that are characteristic of the twelve closet tragedies linked with Mary Sidney reflect a deliberate effort to correct the defects of the public stage referred to in Philip Sidney's *Defense of Poesy*. Computational stylistics offers a way of exploring *Mariam* in the context of other dramatic texts from the period and can also provide new insights into the apparent contradictions that the text embodies.

A cognitive analysis provides a framework in which the variables identified as statistically significant in a computational study can be linked to rhetorical to textual features and connections that can add to critical interpretations associated with more traditional approaches. Burrows-style computational stylistics shows that function word analysis highlights differences as well as points of intersection between diverse early modern dramatic traditions. The combination of Cognitive Grammar and computational stylistics show that although *Mariam* is identified as a closet play, the text reveals a distinctive use of particular auxiliary verbs, modal auxiliaries, and conjunctions. Using Cognitive Grammar, these features in *Mariam* can be analysed as part of a rhetorical strategy that contributes to a reading of the play as an exercise in moral philosophy and an exploration of epistemological uncertainty.

Investigating Minor Textual Differences Using Approximate String Matching Algorithms

Maki Miyake

Osaka University, Japan

Our on-going study aims to investigate the minor textual variants among the modern critical editions of the Greek New Testament. With neither simple word frequency data nor term frequency-inverse document frequency (tf-idf) weighted data, it might be difficult to identify the differences between two documents that are mostly consistent with each other. Indeed, such documents usually have a very strong correlation coefficient calculated by word frequency, which indicates they are similar overall. In contrast, the result of the correlation coefficient based on tf-idf weights changes the perspective completely. The tf-idf weight has widely been used in information retrieval. One of the major characteristics of this measure is that the more widely and repeatedly a term occurs across documents, the less weighting is given to the term. For this reason, it is possible to extract the distinctive words in each document. However, the tf-idf measure does not work in the case of extremely similar documents. The result completely sets aside the common word occurrences of the documents that account for most of the content and focuses too much on the minor individual differences among the documents.

For the purpose of overcoming these problems mentioned above, we made use of more sophisticated as well as easy-to-handle methods to distinguish the differences among the major critical editions of the Greek New Testament. We especially focused on two promising techniques; one is a fuzzy string matching method used in the field of record linkage, and the other is a pairwise sequence alignment algorithm that is also known as a dynamic programming in bioinformatics.

The fuzzy matching technique, which was recently proposed by Wang et al. (2011), consists of three fuzzy-token similarities such as fuzzy Dice/Cosine/Jaccard similarities. For example, the fuzzy cosine measure is combined with the Levenshtein edit-based distance and the cosine token-based similarity. Taking advantage of both the edit-based and token-based distance, this hybrid method enables us to approximate string matching with the higher accuracy.

Pairwise Sequence Alignment was originally used to identify regions of similarity between two sequences of DNA, RNA, or protein in bioinformatics.

Sequence alignments have been also applied to non-biological sequences, such as those in natural language processing. We applied one of the local sequence alignments, called the Smith-Waterman algorithm, which was proposed in 1981. The method could extract the portion of highest similarity between two documents and calculate the differences by aligning the strings more accurately than simply aligning by occurrence order.

We applied these computational measures for comparing texts among the critical editions of the Greek New Testament. Following the consensus that the Nestle-Aland edition is considered to be one of the most reliable critical tools for biblical studies, we set the latest 28th edition as the standard text and compared it with other critical editions such as Westcott-Hort (1881), Scrivener (1894) and Robinson-Pierpont Majority Text (2005). Each similarity result successfully indicated the similarities and the differences among the editions, which was not achievable using the methods with word frequency profile.

Wang, J., Li, G. and Fe, J., (2011), Fast-join: An efficient method for fuzzy token matching based string similarity join, Proceedings of the 2011 IEEE 27th International Conference on Data Engineering, p.458-469.

Smith, T. F. and Waterman, M. (1981). Identification of common molecular subsequences, *Journal of Molecular Biology*, 147, p.195-197.

Narrative Style in the Scientific Romance Mode: an experiment in method and genre analysis.

Naomi Kathleen Fraser

University of Newcastle, Australia

Most multivariate studies that have shown stylistic distinctions between genres are based on form rather than content; however, there is a basis for exploring the linguistic signals of genre through multivariate analysis, as demonstrated by Matthew Jockers (Jockers 63–70). It has been noted that a multivariate analysis of genre is distorted by other signals that are often stronger, such as author and time signals but methods have been explored which can suppress or fix variables (Hoover 350). Before exploring how genres function linguistically, I propose that we must explore how modes function. Jussi Karlgren has suggested that readers observe variations in style according to utility and argues that a computational study of genre should explore features identified by their utility rather than features that are statistically determined (Karlgren 115, 118). Thus, if we are to analyse how modes function within genre, we should initially determine whether there are shared features of narrative style in a particular mode. This paper will analyse narrative style in selected works by Jules Verne and H.G. Wells that belong to a mode of story-telling known as ‘scientific romance’ (Attebery 34). The texts within the ‘scientific romance’ mode share a similar communicative purpose: to frame futuristic narratives that explore rational, scientific or societal alterations and to align with the concerns of society at large through rational scientific explanation (Wolfe 24). Accepting this function of ‘scientific romance’, we can commence analysis of how language operates within these texts to create this particular function of the genre science fiction. A particular focus of this paper is the question of whether there are shared patterns in language in how futuristic worlds and social concerns are framed. Stylistic studies in science fiction and fantasy have shown that there are syntactic devices that can be employed by authors to present alternative worlds and realities (Mandala) and that the suppression of communication is a feature of dystopian narratives (Stockwell). This study will specifically analyse features of language in a selection of

narratives that are futuristic, dystopian and utopian. A study in style variation across the texts will be conducted to identify patterns in the presentation of futuristic features of narrative and the relationship between these trends and content before turning to a descriptive analysis of the relationship between language and the function of the 'scientific romance' mode. This study proposes to employ multivariate methods, specifically cluster analysis and principal component analysis, using David Hoover's modification for style variation studies (Hoover).

Attebery, Brian. *Strategies of Fantasy*. Bloomington: Indiana University Press, 1992. Print.

Hoover, DL. "Multivariate Analysis and the Study of Style Variation." *Literary and Linguistic Computing* 18.4 (2003): 341–360. Web. 16 Apr. 2013.

Jockers, Matthew. *Macroanalysis: Digital Methods and Literary History*. Urbana: University of Illinois Press, 2013. Print.

Karlgren, Jussi. "Textual Stylistic Variation: Choices, Genres and Individuals." *The Structure of Style: Algorithmic Approaches to Understanding Manner and*

Meaning. Ed. Shlomo Argamon, Kevin Burns, & Shlomo Dubnov. Heidelberg: Springer, 2010.

Mandala, Susan. *Language in Science Fiction and Fantasy: The Question of Style*. London: Continuum International Publishing, 2010. Print.

Stockwell, Peter. *The Poetics of Science Fiction*. Harlow: Longman, 2000. Print.

Wolfe, Gary K. *Evaporating Genres: Essays on Fantastic Literature*. Middletown: Wesleyan University Press, 2011.

2.2: Archives & Collections I

Location: University Club Auditorium

Presentations

A Great War story: creating a digital future for the European War

Collecting Project

Maggie Patton

State Library of New South Wales, Australia

The State Library of NSW has an extraordinary WWI collection due largely to the foresight of William Ifould, the Principal Librarian from 1912 to 1942. Ifould recognised the cultural significance of documenting the personal stories of war, and initiated the Library's first significant collecting drive, the *European War Collecting Project*, in 1919. He actively sought out diaries, correspondence, photographs and related material from servicemen. He advertised in newspapers and servicemen's publications with the promise of '*Good prices will be paid for good material*'.

The Library now holds over 1100 volumes of diaries and letters written by over 450 service men and women. As part of the Library's WWI Centenary Program this collection will be digitised, transcribed and released online. Through a dedicated web presence and using a variety of digital platforms the Library plans to engage a range of audiences from local schools and community groups through to the digital humanities network, ensuring that Ifould's European War Collecting Project will generate a legacy of diverse research projects unimaginable to Ifould in 1919.

This paper will describe the World War I project from digitisation through to the development and launch of the Library's Centenary website planned for June 2014. It will discuss the inspiring and challenging aspects of executing Australia's first large scale manuscript digitisation project including rights management issues, metadata creation, the complexities of digitizing non standards formats and online publication. The visual representation of each diary will include full transcriptions of the hand-written entries. These transcriptions will be

generated using the combination of a commercial offshore transcription service and a group of dedicated volunteers utilising a web based transcription interface developed inhouse.

Data generated through the digitisation of the diaries, photographs, maps and related material will be open and accessible for researchers to download, reuse and recreate. Our collection experts will also be curating digital stories and sets of data focusing on areas of particular strength in these collections.

Libraries have traditionally provided the primary resources for research and publication. We plan to ensure that this role continues in a digital future and that primary resources and the foundation documents which cultural institutions acquire and preserve become part of our digital heritage. The WWI Centenary Program aims to provide an exemplar for cultural institutions transforming their analog collections into a digital format freely available online.

Curation, Conversion, Coexistence: Legacy research data, the archive, and the digital humanities

Antonina Lewis, Jennifer Warburton

The University of Melbourne, Australia

This paper will report on issues, interventions, and outcomes of a data curation project undertaken by the authors as part of the 'Immersive Informatics' course at the University of Melbourne - a pilot programme for Research Data Management training, developed in partnership with the University of Bath and offered to participants for the first time in 2013. The current data curation project extends on work already completed by Dr Lewis and others as part of an ARC funded research project that resulted in publication of the online resource *Saulwick Polls and Social Research* (www.saulwick.info).

Our starting point for the project is a circumscribed subset of the Irving Saulwick archive: a legacy research collection of humanities data in the form of focus group interviews recorded onto audiocassettes and VHS videocassettes over a twenty year period (1986-2006), and the digital surrogates created for these materials. Our data curation project reflects on the following questions:

1. What does "curation" mean in the context of research data and the digital humanities?

...considering and acknowledging differences and affinities in how this term is used and understood by different sectors of cultural practitioners and information professionals, we seek to explore tensions and new understandings associated with the terminology

... advance "the archive" as a parallel example of translated terms.

2. How do we conceptualise, enact, and represent relationships between physical artefacts and their digital surrogates?

... archival retention of source artefacts as well as digital preservation and dissemination versions [3.75 TB data created ... choice of formats & metadata schemas].

3. Where do we ethically situate boundaries for access and reuse of a set of legacy records with "grey" or indeterminate consent attached?

... are these boundaries fixed or mutable

... and do they change for “research data” vs “an archival collection”.

Googling the Archive when the Archive has been Googled

Paul William Genoni¹, Tanya Dalziel²

¹Curtin University, Australia; ²University of Western Australia

When Google announced in November 2008 that it would host online one of the world's largest compendium of photographic images thanks to its collaboration with *Life* magazine, it also said something about the state of the archive and humanities scholarship in the digital age. While boasting of the millions of images that would be available for viewing in accordance with the company's audacious “mission to organise all the world's information and make it universally accessible”, the publicity also signalled the limitations of the off-line archive. “Only a small percentage of these images have ever been published,” the press statement read, “The rest have been sitting in dusty archives in the form of negatives, slides, glass plates, etchings and prints.” With the rise of Google, the archive as humanities scholars have known it comes to look inefficient, inactive, indifferent, and inaccessible. The online version, by contrast, promises something ‘better’, more alive, effective and interactive, thereby altering what is meant by, what comprises, and who assumes access to ‘the archive’. Without reinscribing the divide between on-line and off-line archives on which Google's publicity turns, this paper recognises it as a prompt for thinking about the place, value and purpose of the archive in digital society and humanities scholarship. To pursue this investigation, the focus of the paper falls on a small subset of images (1,600 black & white photographs) that constitute part of this huge image repository, and which are also now available from Getty Images' vast commercial ‘bank’.

Life photographer James Burke took the photographs in question during a one-month period on the Greek island of Hydra in 1960. Taken at a critical juncture in the Island's development, Burke's photographs document an expatriate bohemian community that included artists and writers such as Australians Charmian Clift and George Johnston; a young Leonard Cohen; leading Greek modernist painter Nikos Ghikas; Greek-American painter Dimitri Gassoumis; Scandinavian novelists Axel Jensen and Tore Pedersen; and American novelist Gordon Merrick.

Burke's photographs, which appear to have been lodged in the *Time-Life* archive soon after having been taken, were virtually unknown before their appearance in the Google database. As digital images, however, they have emerged as a treasure trove for researchers interested in the interrelationships between visual images and biography, who would otherwise have no knowledge of their existence. They also raise many questions about the nature and use of on-line archives, as well as broader, complementary inquiries about the state of the (digital) archive, its uses and users, which this paper explores. As befits images that appear in ‘stock photography’ service such as Getty Images, the meta-data relating to our ‘case study’ images, which are difficult to recognise as standard ‘stock’ material given their very specific subjects, is minimal and in some cases incorrect. This detail, or rather the lack of it, highlights the role researchers (and in the case of Leonard Cohen, fans) have in constructing, interpreting and narrating such an archive and its constituent items. In effect, as part of research project based on this artist community, we have self-curated a recognisable and meaningful collection from a database that does not itself categorise or recognise the photographs in such a manner, which prompts us to interrogate both the logic at work in these on-line archives and our own practices of scholarship and interpretation.

The paper, then, will both range over a series of questions that arise for humanities scholars from our experience of discovering and working with specific digitised images and digital archives, and reflect on Google's enthusiastic vision of the archive in the digital age.

2.3: BOF: Digital Humanities and the Tyranny of Distance

Location: Seminar Room 2

Presentations

Digital Humanities And The Tyranny of Distance

Erik Malcolm Champion¹, Christof Schöch², Matt Munson³, Toma Tasovac⁴
1MCCA, Humanities, Curtin University, Australia; 2Chair for Digital Philology, University of Würzburg, Germany; 3Göttingen Centre for Digital Humanities (GCDH), Germany; 4Belgrade Centre for Digital Humanities, Serbia

Collaborative writing

This contribution reports on experiences made with collaborative writing in the DARIAH consortium (Digital Research Infrastructure for the Arts and Humanities, www.dariah.eu). DARIAH is a distributed research project involving 12 different European partners; tools supporting various collaborative writing and project coordination tasks have been used over a considerable period of time. Collaboration across geographical distance is essential for this and many other projects, but the existence of conflicting requirements of scholarly collaborative writing processes make a generic solution difficult. Among these requirements are realtime collaborative writing, flexible word-level commenting, footnote support, version control, access rights management, publishing options and open-source availability of the tool itself. Currently available technical solutions do not meet all of these requirements. Tools discussed include Etherpad, Mediawiki/Confluence, GoogleDrive, Dropbox and Wordpress. Finally, one promising solution will be discussed which is still in early stages of development, namely Penflip (www.penflip.com), a GitHub front-end for text composition.

How Can VREs Enhance Distance Research

A recent report on virtual research environments (VREs) from the Joint Information Systems Council in the UK (JISC) found that, even after 6 years of funding and study, "the 'emergent community of practice' has failed to grow significantly beyond the pool of practitioners in direct receipt of JISC project funds,"[1] perhaps it is time to consider whether VREs truly can be a useful addition to humanities research. This paper will discuss the areas in which scholars should expect VREs to assist them in distance research (access to the same tools, data, and workflows in a single environment) and the price they will need to pay for these advantages (either significant time and energy to develop their own environment or being satisfied with a pre-existing solution). VREs can be an excellent tool for distance research, but one for which a significant price must be paid given the current state of existing VRE platforms.

Recognizing Distance: On Multilingualism in Digital Infrastructures

Infrastructures are installations and services that function as "mediating interfaces" or "structures 'in between' that allow things, people and signs to travel across space by means of more or less standardized paths and protocols for conversion or translation." [2]
Infrastructures are in the business of overcoming distance: they have always been seen as motors of change propelling society into a better and brighter future. It is all too tempting —

and all too easy — to approach the question of digital research infrastructures uncritically by embracing the master narratives of efficiency and progress without discussing the larger and more complex implications of institutionalizing networked research. A digital infrastructure is not only a tool that needs to be built: it is also a tool that needs to be understood. In this talk, I will address the challenge of multilingualism in research infrastructures evolving against the backdrop of global capitalism in its electronic mode, the so-called “eEmpire” [3] How can we make sure that digital infrastructures — not only the ones we are trying to build now, for ours are baby steps, but the future ones, the ones we hope to see built one day — do not turn from being power grids into grids of (hegemonic, monolingual, monocultural) power?

The 3D world is your stage

Which conferencing and distributed modeling tools are particularly appropriate to research and collaboration in the spatial and artefactual humanities? This talk will briefly outline needs, issues and promising services and working prototypes.

[1] http://www.immagic.com/eLibrary/ARCHIVES/GENERAL/JISC_UK/J100315M.pdf , p. 21.

[2] Badenoch, A. and A. Fickers (eds.), *Materializing Europe: Transnational Infrastructures and the Project of Europe* (Basingstoke, Hampshire; New York:

Palgrave Macmillan), 11, [3] Raley, R. (2004). eEmpires. *Cultural Critique* 57, 132.

15:20- 16:50 **Sessions**

3.1: Language & Text III

Location: Seminar Room 1

Presentations

Navigating the 18th-Century Republic of Letters: Sequence Alignment as a Method for Scalable Reading in Large Text Corpora

Glenn H Roe

Digital Humanities Hub, Australian National University

The promise of “big data” approaches to humanities research has led in recent years to a host of innovative computational and algorithmic methods for exploring the growing digital human record. These techniques—from data mining to distant reading—offer students and scholars new perspectives on otherwise intractable data sets; perspectives that would have previously been unimaginable. The danger, however, in these kinds of “macro-analyses”, is that scholars find themselves increasingly disconnected from the raw materials of their research, engaging with massive collections of texts in ways that are neither intuitive nor transparent, and that provide few opportunities to apply traditional modes of close reading to these new resources. This paper aims at addressing these concerns through the use of a text mining software package that applies sequence alignment algorithms—borrowed from the field of bio-informatics—to large-scale heterogeneous data sets in order to identify varying degrees of “intertextual” relationships.

The software, named “PAIR: Pairwise Alignment for Intertextual Relations” [Horton et al., 2010], was developed collaboratively by the University of Chicago’s ARTFL Project, and is currently being tested over several large literary and historical data sets in both Chicago and at the University of Oxford. In creating PAIR, we attempted to adapt existing techniques to suit the particular needs of humanities scholarship. Many algorithms such as BLAST [Altschul et al., 1990] exist for identifying duplicated DNA, for example, but these tend to emphasize speed

over completeness. In text analysis, our corpora are small enough and our interest deep enough that we can emphasize retrieving as many hits as possible, while at the same time returning results that are of maximal interest, avoiding a preponderance of banal commonplaces and other formulaic phrases.

To achieve this, the PAIR system first indexes documents by breaking them into overlapping sequences of words, or “n-grams”, and then creates a database of the occurrences of each n-gram in a given corpus. This index allows for the discovery of text reuse by looking for occurrences of the same word sequences shared between documents or within different parts of the same document. Through various tuning parameters, this technique can be made flexible enough to find sequence matches with minor differences in word order, missing words, orthographic variations, misrecognized characters, and other textual variants. The first part of this paper will thus examine the philosophy behind the PAIR approach, the algorithmic design, its implementation as an open-source Perl module, and its application to a variety of tasks relevant to humanities research.

Building upon this methodological base, I will then present my current work using PAIR to explore the circulation of ideas and texts in the 18th-century “Republic of Letters”, and in particular, through the more than 60,000 letters contained in the University of Oxford’s *Electronic Enlightenment* database. The unique epistemological status of correspondence collections allows us to explore a more expansive area of “intertextuality”—one that incorporates recent work on social networks and communication circuits, historical media studies, and, more generally, literary, intellectual, and book history of the Enlightenment period.

By reintroducing this expanded notion of (inter)textuality into algorithmic and data-driven methods, we can perhaps bridge the gap between distant and close readings, through an intermediary mode of scholarship I term “scalable” reading, to borrow a term from Martin Mueller. I will further argue that the digital humanities as a field is perhaps uniquely poised to function as an interdisciplinary locus for just this sort of scalable textual analysis, moving between the distant and the close, the macro and micro, by way of a digitally-assisted “meso-analysis” informed both by new computational methods and the centuries-old tradition of textual scholarship.

How to discover textual groups

Timothy John Finney

Vose Theological Seminary (Research Associate)

Multivariate analysis (MVA) can be applied to the New Testament textual tradition in order to investigate grouping among its witnesses. This article applies certain MVA methods to a number of example data sets. Each method operates on a matrix which tabulates distances between pairs of items in a data set. The simple matching distance, which is the proportion of disagreements, can be used as a metric for calculating distances between New Testament witnesses.

Analysis methods called *classical multidimensional scaling* (CMDS) and *divisive clustering* (DC) are useful for revealing group structure when it is well defined. However, they are less useful when grouping is not very distinct. A method called *partitioning around medoids* (PAM) provides another way to divide a data set into groups. Local maxima in a plot of a statistic called the *mean silhouette width* (MSW) indicate preferred numbers of groups.

Statistical analysis of a data set allows upper and lower critical limits to be defined for the distance between a pair of witnesses. Distances between these limits are not significant in the sense that the same range of distances is expected to occur for generated pairs whose states are randomly chosen from the available pool. Distances that are either less than or greater than these critical limits are not likely to happen by chance. A distance less than the lower critical limit indicates an adjacent relationship while one greater than the upper limit implies an opposite relationship.

Applying CMDS, DC, and PAM analysis to data for the Gospel of Mark reveals interesting features of the textual landscape. Witnesses tend to form groups which have points of contact with conventional categories such as the “Alexandrian,” “Byzantine,” “Western,” and “Eastern” types identified by prior generations of researchers. Multivariate analysis can also be used for novel purposes such as identifying group representatives, group cores, and readings useful for classification purposes.

Interpreting the text mine – a cautionary note from (historical) corpus linguistics

Susanne Flach

Freie Universität Berlin, Germany

While researchers in the philologies have been using historical corpora for a long time and for what they represent — a reflection of language change, cultural topics, societal ideas and philosophical currents of their times (and text basis) — scholars from other fields, such as cultural psychology, have recently begun to mine large-scale historical texts in search for answers of the change of society.

The ease of access to tools (such as the Google NGram Viewer) often means that such scholarship is undertaken with an unfortunate neglect of the nature of corpora, text basis and, most problematic perhaps, the nature of language and language change. Studies in cultural psychology have repeatedly inferred straightforward evidence of cultural change from simple frequency counts (Greenfield, 2013; Kesebir & Kesebir, 2012; Twenge, Campbell, & Gentile, 2012). However, such studies, going by the name of “culturomics” (Michel et al., 2011), have consistently been flawed both methodologically and theoretically.

In this paper, I will present two case studies illustrating the flaws of such studies from the point of (historical) corpus linguistics. Case study one will deal with the shortcomings on the level of the lexicon (“individualistic” vs. “communal” words), case study two with the structural, syntactico-grammatical level (“obligation” vs. “choice” constructions). Both case studies shall demonstrate how many such investigations ultimately lead to naïve interpretations about data to arrive at rather bold conclusions of the nature and change of society. From a theoretical, methodological and empirical perspective, the guiding issues are:

- The polyfunctionality of language structure and the inevitability of language change
- The problematic attempt to map complexity to simplicity to complexity
- The problems of quantitative inference of an equally qualitative phenomenon

As the field of corpus linguistics has so far remained surprisingly silent with respect to the claims in “culturomics”, this paper shall attempt to make such a contribution. It will be of elevated interest to anyone working in areas that tap into the structural and semantic analysis of language; it shall provide cautionary notes and impulses for methodological refinement in digital historical linguistics.

Greenfield, P. M. 2013, The changing psychology of culture from 1800 through 2000. *Psychological Science*, Published online August 7, pp. 1–10.

Kesebir, P., & Kesebir, S. 2012, The cultural salience of moral character and virtue declined in twentieth century America. *The Journal of Positive Psychology*, 7(6), pp. 471–480.

Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., ... Aiden, E. L. 2011, Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014), pp. 176–82.

Twenge, J. M., Campbell, W. K., & Gentile, B. 2012, Increases in individualistic words and phrases in American books, 1960–2008. *PLoS ONE*, 7(7), pp. e40181.

15:20- 16:50 **Sessions**

3.2: Archives & Collections II

Location: **University Club Auditorium**

Presentations

Deploying Ontologies in the Humanities and Creative Arts

Toby Burrows¹, Deb Verhoeven²

¹University of Western Australia, Australia; ²Deakin University, Australia

This paper will examine issues related to the deployment of ontologies in digital infrastructure for the humanities and creative arts.

Ontologies – for the purposes of digital information systems – are formal, structured representations of knowledge within a domain, modelled as the types of concepts or objects which exist within that domain, together with their properties and relationships. They are normally distinguished from vocabularies, which lack the hierarchies of classes and sub-classes usually associated with ontologies.

Ontologies are widely used in knowledge management systems and are critical to automated reasoning. This is especially the case in the context of faceted searching and browsing, but also applies to recommendation systems and to business systems involving machine-to-machine negotiations. Ontologies are closely connected to the Semantic Web and to its implementation through Linked Data technologies.

Ontologies are most applicable in knowledge domains with an agreed common vocabulary and a shared understanding of the semantic and conceptual structures of the domain. A typical example from the sciences is the Gene Ontology, which is being developed to unify the representation of genes and gene products across all species.

Processes for identifying, describing and categorizing phenomena – both individuals and concepts – are central to research in the humanities and the creative arts. Because of this, ontologies are potentially very valuable. But interpretation and framing are also central to humanities research and analysis, and there are inherent differences in the perspectives of researchers: social, cultural, political, geographical, and disciplinary. From this point of view, ontologies pose a range of problems. These include:

- Variations in terminology: The same concept or phenomenon may be described using different terms by different researchers, let alone by the wider community.
- Vagueness of terminology: The same term may be capable of referring to different concepts or phenomena, depending on the context.

- Historical change: The understanding of a knowledge domain is likely to have changed dramatically over time, along with the vocabulary and values used. Ontologies do not usually have a temporal dimension.
- Multilingualism: Much humanities research involves languages other than English, either for the subject of the research or for the research discourse itself. Most ontologies, on the other hand, use a standardized (and often quite formal) version of the English language.
- Interdisciplinarity: There is an obvious tension between (and even within) humanities disciplines in their terminology, methodologies and intellectual models. Most upper-level interdisciplinary ontologies are very generalized and do not adequately reflect the complexities of particular domains.

Does this mean that ontologies are inapplicable across the humanities and creative arts? Are vocabularies the most we can aim for? Do we have to fall back on purely linguistic associations discovered through keyword and phrase searches? Is it inherently impractical to build knowledge management systems of the kind developed for scientific disciplines? To what extent can technologies which enable 'socially linked' data address these questions? This paper will examine these questions, drawing on experiences gained during the Humanities Networked Infrastructure (HuNI) project. It will offer some suggestions for alternative approaches to the use of ontology-like approaches in the humanities and creative arts, aimed at building a network of ontological and semantic assertions which can embody a range of different perspectives on knowledge and lay the foundation for future research and development.

Extracting Relationships from an Online Digital Archive about Post-War Queensland Architecture

Jane Hunter¹, John Macarthur¹, Deborah van der Plaats¹, Janina Gosseye¹, Andrae Muys¹, Craig McNamara¹, Gavin Bannerman²

¹The University of Queensland, Australia; ²The State Library of Queensland

The "Architectural Practice in Post-War Queensland: Building and Interpreting an Oral History Archive" project is a collaboration between the University of Queensland, the State Library of Queensland (SLQ) and four of the longest-standing architectural firms in Queensland. The project's aim is to build a comprehensive online multimedia digital archive that documents architectural practice in post-war Queensland (1945-1975) – a period that was highly significant but largely undocumented. The goal is to use innovative Semantic Web technologies to link tacit knowledge extracted from individual oral histories to tangible knowledge (drawings, books, photographs, manuscripts) that exists within personal and firm archives as well as State and institutional archives and libraries.

The methodology involved firstly conducting and recording a series of oral history interviews and public forums with the key architects from this period. These events comprise both private interviews, one-on-one conversations between the project team and architect/s as well as a number of larger public forums held at the SLQ and dedicated to a specific theme (education, style, climate, region, etc.).

The oral history interviews and the public forums are recorded, captured as digital files (.wav and .avi) and transcribed. Text processing is applied to the transcripts to semantically tag key entities mentioned in the interviews and extract new knowledge in the form of RDF graphs.

The resulting RDF graphs document relationships between architects, firms and buildings (with attribution to the source) and are able to be displayed, edited, saved and re-used via the LORE compound object authoring software.

This paper will describe our approach to establishing the online archive and evolving knowledge-base. The system uses the Omeka content management system to support the upload and description of content (oral history files, transcripts, photos, drawings, articles etc.) by the collaborators. In addition the system provides the following components and functionality:

- An OWL ontology defines the core classes, class hierarchies and properties/relationships associated with each class e.g., architects, firms, buildings, documents;
- D2RQ is used to convert Omeka metadata to RDF and save it to a Sesame RDF triple store with a SPARQL query interface;
- A Euler reasoner is applied to the Sesame RDF triple store to reconcile common entities (via URLs) and to infer relationships between key entities (architects, firms, structures/buildings and documents);
- A Search and Browse engine (based on solr) enables users to search for specific entities or perform full-text searching across all transcripts and articles (and jump to matching audio/video segments);
- Word clouds and word frequency histograms have been generated from the oral history transcripts using D3;
- User-authenticated annotation tools enable users to semantically tag transcripts by identifying people, places, buildings, firms, events mentioned in the interviews;
- Mapping and timeline interfaces enables users to interactively retrieve information (interviews, photos, drawings) about buildings, people or events via maps and timelines;
- The LORE tool enables the visualization, editing, sharing and re-use of RDF graphs that document relationships between architects, firms, buildings, and related documents.

Architects who studied and worked in Queensland during the post-war period are also invited to register, login and submit their own details including a chronology of practice and to provide feedback to the existing content. An additional blog monitored by the project team encourages the broader community (those outside the profession) to comment on aspects of post-war architecture (e.g., nominate their favourite building) and to upload related materials such as photographs or plans.

Finally, our paper will describe the challenges that this multi-disciplinary project faces including: attracting and retaining an active community of contributors; ensuring the archive's sustainability, resolving issues of identity and implementing quality control over the community-generated content.

FloraCultures: The Nature of a Biodiversity Archive for Plant-Based Cultural Heritage

John Charles Ryan

Edith Cowan University, Australia

This paper examines the use of digital archival technologies for the conservation of plant-based cultural heritage in Perth, Western Australia. Whereas digital archiving approaches have been applied to the conservation of cultural heritage (for example, MacDonald, 2012), relatively little has been published on the use of digital technologies for natural heritage conservation and promotion. On a conceptual level, what constitutes natural heritage is contested—often bifurcated from cultural heritage and overly narrowed through an exclusive emphasis on material artifacts. However, recent developments in intangible natural heritage (for example, Dorfman, 2011) challenge clear divisions between cultural and natural—as well as tangible and intangible—forms of heritage. These developments impact the archiving of heritage in all its forms.

The aim of FloraCultures (www.FloraCultures.org.au) is to conserve the heritage (defined broadly and inclusively) of Perth's plants through a combination of archival, ethnographic and design strategies. The resulting web archive will feature a broadly conceived collection of heritage information—from recorded interviews with long-term conservationists to items of literary, artistic, and historical interest that in some manner originate from (or comment on) the indigenous flora of the city. This widely dispersed group of materials constitutes the plant-based cultural heritage gathered within the FloraCultures web archive. A city of remarkable botanical diversity, Perth also faces significant plant conservation challenges, including urban and suburban development and plant diseases. As plants vanish so do the irreplaceable heritage values associated with them. In particular, the paper evaluates the role of the archive in the preservation of biodiversity heritage. In *Under Suspicion: A Phenomenology of Media*, Boris Groys outlines a way of thinking about the fundamental place of the archive in culture. Furthermore, Rick Prelinger characterises the 'accessible archive' as necessitating a community of users, whereas Vivian van Saaze advocates the principles of change, intervention and production in archive management. In describing the role of the archive broadly in Western history, culture and society, Groys proposes the term 'cultural economy' as 'the exchange that takes place between the archive of cultural values and the profane space outside this archive' (Groys, 2012, p. 1).

By definition, an archive preserves immaterial (digital) or material (artefactual) items of cultural importance. Conversely, it excludes irrelevant or unvaluable things, relegating these items to the 'profane' world that exists outside of the archive's ambit. Most importantly, the boundaries of the cultural archive are fluid and changing. Things in the profane space become valued and those in the archive become irrelevant or ignored over time (Groys, 2012, pp. 1-2). For Groys, the determination of what is significant and, therefore, to be included in an archive (what he terms the 'New') and what is irrelevant and, therefore, to be excluded (what he calls the 'Old' or 'noncollected reality') is a matter of power and politics. Through the FloraCultures case study, the paper will conclude with prospects for cultural and natural heritage conservation through emerging digital archiving theory and practice.

Dorfman, E. (2011). *Intangible natural heritage: New perspectives on natural objects*. Hoboken: Taylor & Francis.

Groys, B. (2012). *Under suspicion: A phenomenology of media*. New York: Columbia University Press.

MacDonald, L. (2012). *Digital heritage*. Hoboken: Taylor & Francis.

3.3: BOF: The Altruistic Crowd – The Ethics of Using Social Data in Humanist Research?

Location: Seminar Room 2

Presentations

The altruistic crowd – the ethics of using social data in humanist research?

Jennifer A Harrison¹, Darren Gibson²

¹UWA, Australia; ²ECU, Australia

The altruistic crowd – the ethics of using social data in humanist research?

Background: “Crowdsourcing” – gathering information from a large number of people (mostly unknown and from online media) is reputed to have first been defined by Howe, in Wired Magazine as early as 2006[1]. Since that time, the public providing information freely has been growing mainly as a consequence of the ever-increasing number of social media tools available to everyone with a mobile device and an Internet connection. In the era of social media, it could be argued that everyone with a mobile device both consumes as well as provides information regularly.

Researchers have not been blind to this phenomenon, and in some disciplines are turning to social media, using tools to capture information, and mining data provided by social media for their research studies. Are there effective frameworks in place to prevent “participants” (willing or otherwise) from being exploited? It could be argued that not every project is suitable for crowdsourcing, nor should social media data being provided in the height of an emergency, often under stressful situations and circumstances be used for research. Often such information may have inaccuracies, is provided with zero control and with limited regard of how this information is used, analysed, aggregated and re-used.

Are the existing research and governance ethics frameworks, still meaningful for data collected, and shared via social media to ensure investigations are altruistic?

Overview: This Birds of a feather (BoF) session will focus on the ethics of data collected using social media tools. The BoF will examine the issues associated with this type of research and its ethics as well as debating if true philanthropic endeavors can remain at the heart of every investigation with data collected from social media.

Outline for session:

1. Overview of crowdsourcing and social data in humanist research (5 minutes)
2. The ethics of research data use (10 minutes)
3. Altruistic use of social media data in research – maintaining a humanist approach? (10 minutes)
4. Discussion with audience (30 minutes)
5. Wrap up (5 minutes)

[1] http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html

17:00 Poster - Visualisation evening reception

Location: John Curtin Gallery

Buses depart from the University Club for the John Curtin Gallery, and returning after the event

Main Conference - Thursday, 20th March 2014

8:45 – 9:15 Registration

9:15 – 10:30 **Keynote - Dr Sarah Kenderdine**

Special Projects, Museum Victoria, Australia (2003–) and is Visiting A/Prof, Director of Centre for Innovation in Galleries, Libraries, Archives and Museums (iGLAM) and Director of Research at the Applied Laboratory for Interactive Visualization and Embodiment (ALiVE), Co-Director of the LUXLAB, City University, Hong Kong.

In widely exhibited installation works, she has amalgamated cultural heritage with new media art practice, especially in the realms of interactive cinema, augmented reality and embodied narrative. Sarah holds the position of Professor, National Institute for Experimental Arts (NIEA), COFA, University of New South Wales (2013–) and Special Projects and, Special Projects Museum Victoria, Australia (2003–). She is Adjunct Prof at RMIT, Melbourne and at CityU Hong Kong, as Director of Research at the Applied Laboratory for Interactive Visualization and Embodiment (ALiVE).

Recent books include PLACE-Hampi: Inhabiting the Panoramic Imaginary of Vijayanagara, Heidelberg: Kehrer Verlag, 2013 and the co-edited, co-authored Theorizing Digital Cultural Heritage: a critical discourse, Cambridge: MIT Press, 2007 (third reprint 2010). In 2013, Kenderdine received an ICOM (Australia) Award for the Museum at Kaladham built in Karnataka, India; the Australian Arts in Asia Awards Innovation Award 2013 for PLACE-Hampi and the Museum at Kaladham; a Digital Heritage International Congress & IMÉRA Foundation (Aix-Marseille University) Fellowship 2013 and; the Tartessos Prize for contributions to virtual archaeology worldwide. For recent work see ALiVE <<http://alive.scm.cityu.edu.hk/projects/>>

Location: University Club Auditorium

Chair: Professor Hugh Craig, University of Newcastle, Australia

11:00- 12.30 **Sessions**

4.1: Digital History & Virtual Heritage I

Location: University Club Auditorium

Presentations

Out of the archives and onto the streets: Creating digital and social experiences with heritage datasets

Darren Francis Peacock¹, Jill Allison MacKenzie²

¹Sociable Technology, Australia; ²University of Adelaide, Australia

The practice of documenting places of cultural and historical significance dates to ancient times. It is one of the fundamental practices in human place making. Yet it generally was not until modern times that the systematic documentation of 'historical monuments and sites' gathered momentum, in Europe from the mid nineteenth century.

In the United States, the Heritage Documentation Program of the National Parks Service was inaugurated with the Historic American Building Survey commenced in 1933. In Australia, community organisations such as the National Trust began surveys of sites deemed to be of cultural, environmental and architectural significance from the 1950's. These surveys evolved into lists of 'classified' sites in each state. From the 1970s, state, local and Commonwealth governments also created and managed increasingly complex datasets of information about heritage sites- encompassing thousands of built, cultural and natural heritage sites across Australia.

Since 1955, the National Trust of South Australia has taken responsibility for identifying, documenting, maintaining, preserving and promoting sites of heritage significance. Along with private individuals, specialist groups, local councils, and the State and Commonwealth governments, an enormous amount of research has been undertaken to identify, document and preserve these important places. This substantial body of research and documentation has generally only been accessible through limited online databases or through physical visitation to research rooms and archives. Furthermore, these datasets are not well integrated with each other or with other existing and related source materials, such as maps, plans, photographs, oral histories, film and video.

This data is an enormously rich source of information about what has been valued in our built and natural environment by successive generations. For the most part, heritage datasets have been created and used to support the technical and administrative processes of heritage management such as legislative requirements for identifying and protecting places of heritage significance. Yet these datasets hold enormous, largely untapped potential as enablers of digitally-mediated social, cultural and learning experiences that engender active participation in exploring, researching and valuing heritage spaces.

Only now is the digital revolution starting to transform how these datasets are accessed, used and (re)created. Our paper discusses and demonstrates how these resources can now be reimagined and reinvigorated through the application of contemporary digital technologies such as social, mobile and geo-spatial and emerging curatorial practices such as open licensing, crowdsourcing and co-creation. The transition from archival artefact to live, spatial and social experiences marks the beginning of a new era in awareness, appreciation and participation in the research, preservation and enjoyment of our shared heritage. Traditional forms of place-based heritage signification- plaques and other interpretive signage- are now increasingly inadequate to satisfy the needs and interests of ubiquitously connected digital citizens. In our work with heritage datasets in South Australia we have developed a range of new digital productswebsites and mobile apps- that repurpose heritage datasets in ways which reconnect this data with the places they describe in situ, in real time.

By re-contextualising heritage datasets within the physical spaces they describe we aim to create digital artefacts and experiences that maximise the encyclopedic, spatial, procedural and participatory properties described by Murray (1997, 2012) as the key 'affordances' of the digital medium.

This 'here and now heritage' made possible by digital accessibility and connectivity enables new ways to engage with and re-read both past and present through heritage sites. Our conclusions explore the implications of these products and experiences for the future management and use of heritage datasets in research and community activation projects.

The social and technological challenges of augmenting the ABC's archival collection: Creating cultural objects

Jonathon Hutchinson
University of Sydney, Australia

During 2011, the now defunct ABC Pool project developed an experiment that sought to combine emerging augmentation technologies with the vast archival collection of the Australian Broadcasting Corporation (ABC). The MyBurb project sought to enhance Australian suburbs by augmenting historical artefacts (ABC archives) across the contemporary suburban space to enrich the 'lived-in' experience for Australian citizens. The project was successfully implemented, but was not widely received by the proposed audience it sought to engage. The relatively low penetration rate reveals a disconnect between the remit of public service media (PSM) to engage its audiences in new and innovative ways and the user activities it seeks to involve. This observation supports the cultural production gap Hesmondhalgh (2007) identified between the production and consumption of cultural goods, which I would argue could be united through technological intermediation. However, technological intermediation is only part of the process, which if coupled with the work of the cultural intermediary (Negus 2010) represents cultural intermediation. How then could cultural intermediation, which is the combination of human or non-human actors facilitating the collaborative production of cultural goods, been implemented to avoid the disconnection between the MyBurb project and the ABC audience? These data presented within this paper emerge from three years of research at ABC Pool where I was embedded as the community manager/researcher in residence.

Design Exchange: promoting collaborative curating of Design Archives data.

Catherine Moriarty¹, Harriet Edquist²
¹University of Brighton, United Kingdom; ²RMIT, Australia

This paper describes the collaborative relationship between the RMIT Design Archive Melbourne and the University of Brighton Design Archives in the UK. Building on discussions relating to geospatial data that took place in 2010 and more recently on collaborative curating, the Design Exchange project was established in 2013 as a network for design archives and museums to explore the sharing of digital data and assets and the development of curatorial projects and expertise across analogue and digital domains.

The skills of architecture and design museum curators have a great deal to contribute to wider discussions about the arrangement of things and how meaning is constructed through their configuration. Design archives, as specialist resources with particular qualities, components and potential for connectivity, bring to the fore important questions about the relationship between the analogue and the digital in research processes. As John Ridener identifies in his study of archival theory, tensions between objective and subjective approaches to the archive have underpinned the formation of archival paradigms over time[1]. Digital environments exacerbate this tension by requiring precise data structures in production yet enabling interrogation released from the confines of hierarchical structures. Indeed, as the digital surrogate and the information architectures it inhabits grow in sophistication and ubiquity, the design archive presents an especially interesting nexus to explore the ramifications on research and interpretative practices.

The aim of the Design Exchange is to engage collaboratively and collectively on research questions that include:

1. How can archive and museum professionals be supported in devising curatorial projects characterised by risk taking, experimentation and with a concern for enhancing public literacy about architecture and design in both real and virtual realms?
2. How can we construct 'exchanges' whereby archive and museum professionals could swap places, or digitised data, to interrogate collections other than their own and to produce new curatorial projects?
3. How can we produce digital footprints that link and embed objects, experts, and institutions, in order to connect them to new audiences?

The paper will present the findings of the first Design Exchange symposium held in Barcelona in September 2013, and discuss the collaborative programme devised as a result. In addition to the details of the project plan and the issues of exchanging digital data about design, this paper will explore the nature of this trans-institutional relationship and how the curating of design lends itself especially to the navigation, rearrangement and reconfiguration of digital resources. The paper will consider the role of digital environments and the implications for scholarship as research behaviours in the humanities begin to acquire a hybrid characteristic with both analogue and digital elements. As such, Design Exchange offers a new model for digital humanities research that crosses academic, professional and public spheres.

[1] Ridener, John. *From Polders to Postmodernism: A Concise History of Archival Theory*. Duluth, MN: Litwin, 2009.

Historian to Hacker in 48 hours: Why do so few humanists attend civic hacking events?

Jo Hawkins

University of Western Australia, Australia

Theme: BUILDING the DH COMMUNITY and PRESENCE / Institutionalisation, interdisciplinarity and collaboration

The past few years have seen a surge in government sanctioned civic hacking events, including the 'National Day of Hacking' in the USA, 'National Hack the Government Day' in the UK and 'GovHack' in Australia. Propelled by a growing Open Data movement, these kinds of events attract problem solvers with a social conscience, with the aim of creating web applications that release the social value of government datasets. Yet, the vast majority of participants are developers, designers and entrepreneurs. Why do so few humanists attend civic hacking events?

In May 2013 I attended Australia's largest civic hacking event, GovHack, where I worked alongside a team of local web developers and designers to build an iPhone app in 48 hours. My paper will outline the development of the app to argue that humanists have much to offer at GovHack, and just as much to learn.

Our 'hack' was conceptualised as a response to a recent groundswell of interest in the local history and heritage of Perth, Western Australia. This trend has been exemplified by the 'Lost Perth' Facebook page, which amassed an astonishing 50,000 'likes' in the three weeks, after launching in May 2013. The 'Lost Perth' community use social media to share and comment on photographs of built heritage and cultural ephemera. The growing photographic archive is uncategorised, difficult to search and lacks historical data and image attribution - resulting in a collection as ephemeral as the forgotten heritage it aims to preserve.

The Pixtory iPhone app was created with the intention of harnessing public interest to increase engagement with the State Library of Western Australia's photographic archive, which contains over 70,000 online images. The app allows users to have an immersive experience with Perth's built heritage, by collecting and exposing geo-tagged historic photos using the Trove API. Our team won two national prizes at GovHack 2013, 'Most innovative transport project' and 'Trove Prize', in addition to the local 'Spirit of GovHack' Award. The project also attracted local and national coverage in press, radio and social media. Drawing from my experience at GovHack, I shall argue that humanists can make important contributions to civic hacking events as storytellers and strategists, ensuring concepts and executions are grounded in real life research problems. Further to this, collaboration with industry brings design and usability to the forefront and provides an opportunity for historians to engage with the commercial logics that drive the competitive world of online start-ups. For researchers new to digital humanities, these outcome-driven events offer a potential gateway into the field, providing opportunities to gain practical experience and establish networks with like-minded practitioners and organisations outside the academy. Digital Humanist, Alan Liu has argued that researchers in the humanities need to work harder to engage with public audiences and articulate the value of what they do. Civic Hacking events provide dynamic spaces for collaboration, experimentation and disciplinary outreach.

Australia <http://www.govhack.org/>

UK <http://nationalhackthegovernment.wordpress.com/>

USA <http://hackforchange.org/>

4.2: Network Analysis & Discovery

Location: Seminar Room 1

Presentations

Finding Mobile Internet Policy Actors in Big Data: Methodological Concerns in Social Network Analysis

Gerrard Goggin, Timothy Dwyer, Fiona Martin, Jonathon Paul Hutchinson
University of Sydney, Australia

Scholarly interest in data privacy and the regulation of mobile Internet has intensified in recent years, particularly following Edward Snowden's 2013 revelations about Prism, the US government's secret communications surveillance and data mining project. Much analysis has focused on the politics and architectures of data privacy regulation and network access. However the surveillance moment also invites scrutiny of academic data gathering and mining online. In open governance movements such as Occupy there has already been considerable debate about the ethics of big data research, particularly where the aim is to track individuals' online agency around political processes and policy activism. With that context in mind, this paper examines the methodological implications of conducting large-scale social network analysis using Twitter for mobile Internet policy research.

Mobile internet is emerging at the intersection of broadband internet, mobile telephony, digital television and new media locative and sensing technologies. The policy issues around the development of this complex ecology include debates about spectrum allocation and network development, content production and code generation, and the design and the operation of media and telecommunications technologies. However not all of these discussions occur in formal regulatory settings such as International Telecommunications Union or World information Summit meetings, and not all are between traditional policy

actors. Increasingly social media platforms such as Twitter and Linked-in host new networks of expertise, informal multi-actor conversations about the future of mobile Internet that have the potential to influence formal policy processes, as occurred during the January 2012 SOPA/PIPA campaigns in the US.

As part of the three year Australian Research Council Discovery project Moving Media: Mobile internet and new policy modes, this research team is mapping and interpreting the interplay between these diverse policy actors in three areas of accelerating media development: digital news, mobile health and locative media. However research into informal policy networks and processes online presents interesting problems of scale, focus and interpretation, given the increased affordances for citizen participation within the international political arenas of social media.

To better understand who these online stakeholders might be in the mobile health field, and how they operate in relation to the normative policy and regulatory circuits, we have adopted a social network analysis methodology, in order to track Twitter-based social relationships and debates. Using a series of hashtags, including #mhealth, #mobilehealth and #healthapps to track ongoing policy-related exchanges, we have begun to identify who is influential in these spaces, what they are talking about and how their input to debate may impact on mobile internet regulation.

This paper will outline that SNA approach and highlight some of the procedural and ethical concerns surrounding big data collection and analysis, which are consistent across contemporary digital humanities research. These concerns include how we can use big data harvesting and analysis tools to align quantitative with qualitative methods, how we can justify our research claims via these tools and how we might better understand and implement these innovative research methods within the academy. In particular the paper will interrogate the methodological suggestion that qualitative methods lead quantitative research, considering instead whether a more rigorous approach is to invert the quantitative/qualitative relationship.

Menzies' book club: Using a relational database to reconstruct the social networks of Robert Gordon Menzies

Caitlin Stone, Jim Berryman

The University of Melbourne, Australia

The University of Melbourne Library holds the personal library of Sir Robert Gordon Menzies, prime minister of Australia from 1939 to 1941 and 1949 to 1966. The collection reflects a lifetime of reading and comprises almost 4000 books (as well as magazines, photograph albums and ephemera) that belonged to Menzies. As well as reflecting the owner's tastes and interests, the collection sheds light on Menzies' extensive political, social and personal networks. We demonstrate in this paper how is it possible to use a relational database to piece together an impression of Menzies' intricate social world, based on the evidence of his books.

Much of the value of the collection lies in the annotations and dedications inside the books. We have discovered that almost forty percent of the volumes bear inscriptions ranging from signatures and brief notes to more personal dedications. These reveal the identity of the giver of the book and often its date of acquisition. A smaller number contain signs of previous ownership, suggesting they were donated, exchanged, or perhaps lent to Menzies (and never

returned). Presents to Menzies from family members and friends are common. The collection also contains numerous books presented to Menzies by international dignitaries, not to mention unsolicited gifts from his admirers and constituents. The exchange of books was a form of 'social networking', and these inscriptions reveal a complex network of associations both within and outside the world of politics.

This paper will focus on our efforts to map the connections between books, writers of inscriptions, events and organisations relevant to the life and times of Robert Menzies. Moving beyond the bibliographic approach, our project incorporates archival research and digital technologies to learn more about Menzies' public and private self. Looking ahead, there is further scope for this project to be used as the basis of a detailed social network analysis.

Personal libraries occupy an ambiguous position in traditional library and archival practice. The professional literature on the management and conservation of personal libraries is therefore limited. In disciplines such as history and literature, however, personal libraries have been used as sources of unique biographical evidence. We are not aware of a comparable project using digital technology to build a social and biographical profile of an individual based on their personal library.

We hope the outcomes of this project will not be confined to the interests of Australian and political history. This new approach to understanding personal libraries will encourage others working in digital humanities to explore these unique collections. Held in a number of Australian academic and state libraries, personal libraries offer great potential for digital applications. In addition to network analyses, these collections provide opportunities for book digitisation and OCR technologies, as well as database design.

This project will be of interest to researchers in book history, digitisation, textual analysis and digital curation. It will also interest technologists, especially database designers and developers.

The online resource resulting from this project, The Robert Menzies Collection: A Living Library, has been published as a work in progress:
<http://www.menziescollection.esrc.unimelb.edu.au>.

Trailblazing Women in the Law: Social Network Analysis and Oral History

Louise Baker

The Australian National University, Australia

This paper will explore the 'Trailblazing Women in the Law' project and attempt to outline the various methodologies used in capturing the life stories of Australian Women Lawyers. This paper questions whether the networks of associations captured in the social network map can predict the likely outcomes of individual choices and experiences, historical policy phases and the future impact of proposed policy developments for women lawyers. In contrast to the previous use of women lawyer's narratives in Australia, this project takes a post-structural approach to oral history, with the intention of both using the explicit content of these women's stories to fill knowledge gaps, as well as considering the internal features of the text, such as omissions and framing. We seek to understand trailblazing and to supplement the existing research on gender in the legal profession via the project's three main analytical emphases; firstly, the interplay of professional and social background in pioneers' hurdles and success, secondly, the embeddedness of trailblazers in networks of connection and finally, the self-conceptualisation and activity of women lawyers as active citizens. Fundamentally, our

work will form a new empirical basis for theorising the value of oral history to legal research. As a part of my research I will be using the Online Heritage Resource Manager (OHRM) technology, which will provide a method of collecting, analysing and displaying the connections, divergences and consistencies across the interview pool. The specific application of social network analysis to the Trailblazing Project will contribute to the evolution of this technology in its own right with the project designated as a case study for testing the application of social network mapping and OHRM to a data set. I am interested in how, and to what end, trailblazing women lawyers are bound up in mutually beneficial or professionally detrimental relationships. The social network analysis method will enable the research team to trace and explore the connections between particular individuals in the legal community and to empirically identify and represent collective trends and potential 'sites' of trailblazing across the diverse backgrounds and decades of practice captured in the interviews. The significance of particular kinds of connection, for example; political or familial, will also be explored. This work will advance both the available research sources and analytical treatment of the impact of networks in professional cultures.

Metaphors for discovery: a visual survey of library and other collection interfaces

Georgina Hibberd

University of Technology, Sydney, Australia

This paper will present a visual survey of metaphors used in the graphic user interfaces (GUIs) of libraries and similar organisations. The intent of this survey is to better understand how these metaphors shape the user's relationship to a collection. My research takes as a starting point Johanna Drucker's (2011) call for a 'humanities approach' to interface by analysing the GUIs of large collections from a visual communication design perspective.

Drucker points out an underlying assumption of current theories that interfaces are 'neutral', consideration of metaphors used within them and the assumptions that underlie those metaphors is rare. Interfaces are viewed as pragmatic and instrumental rather than rhetorical and persuasive.

Within library catalog interfaces metaphors are often tied to the physicality of books and to classification systems used by the library. Now that many libraries are holding digital editions of books, or removing public access to book holdings - for instance, the underground automated storage and retrieval system at the University of Technology, Sydney - there is room to question the interface metaphors of collections. There is also the potential to explore the way in which interface metaphors can be viewed through a humanities lens.

David Weinberger (Weinberger 2007) explains how the way in which we impose order and thus access to physical objects and information has been freed by the digital. In Weinberger's three "orders of order" we have moved from a "first order" limited by matter, through a "second order" where information about objects is maintained separate from the objects themselves through to the "third order" which is miscellaneous. Order is defined dynamically, through searches, selections and recommendation engines. In an automated storage system, where books are stored by size and retrieved by robot, the catalog doesn't need to relate to the storage system because patrons don't have to locate the book itself. This provides us with the potential to explore what could be designed when you break the relationship of the Dewey system to the physical arrangement of books when users no longer need to understand their physical location. There is an opportunity to explore the potential of new and novel

metaphors to both represent the collection and to aid in experiencing and understanding the collection. Due to the size of digital collections, interfaces for search have focused on narrowing down options. Some library users use the physical shelves as a vehicle for serendipitous discovery, after performing an initial search online. Removing physical access to some or all of the collection could reduce the potential for serendipitous discovery. My research aims to explore the potential for metaphor to support serendipitous discovery within digital collections using information retrieval models that are drawn from disciplines outside of library and information science. For example, Marian Dörk (Dörk, Carpendale et al. 2011) and his colleagues propose the figure of the Flâneur to conceptualise an information-seeking behaviour that is meandering, whilst effective. The information flâneur finds pleasure in information-seeking; s/he opens up the possibility of an engagement with a library collection that is pleasurable, poetic and serendipitous. Models such as the flâneur could offer us ways in which to renegotiate our relationship with collections whose materiality is constantly changing. This paper presents doctoral work in progress.

Dörk, M., et al. (2011). The information flâneur: a fresh look at information seeking. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM.

Drucker, J. (2011). "Humanities Approaches to Interface Theory." Culture Machine 12: 1-20.

Weinberger, D. (2007). Everything is miscellaneous: The power of the new digital disorder, Macmillan.

4.3: BOF: Building DH Community on the Regional Scale - Notes Towards a Multimodal Regional Digital Cultures Collaboratory

Location: Seminar Room 2

Presentations

Building DH Community on the Regional Scale: Notes Towards a Multimodal Regional Digital Cultures Collaboratory

Daniel Powell

University of Victoria, Canada

Writing in 2004, Cathy Davidson and David Theo Golberg argue that the “new humanities” “require new structures . . . new models for researchers to work across disciplinary boundaries, making use of databases and resources that no one scholar, or department, can maintain. That requires planning at an institutional level.” This presentation will outline the theoretical contexts of such planning at the multi-institutional, regional scale, integrating theory with practitioner interviews from a variety of institutional contexts, followed by group discussion. Ideally, this BofF session will be grounded in a discussion of regional DH development in the Cascadian context and quickly pivot to the relevance of similar infrastructure in the Australasian environment. If the Program Committee would prefer this as a long paper presentation, similar ideas will be presented but without the extensive discussion a BofF would allow. This proposal directly addresses conference them five (Building the DH Community and Presence) and six (Indigenous and Cross-Cultural Digital Research).

Primarily based on the regional coherence of the Cascadian Bioregion (the area of Canada and the United States including southern Alaska, British Columbia, Washington, Oregon, and northern California), this presentation will suggest that one possible scenario for “glimpsing the digital future” of humanistic inquiry is a “thickening” of collaboratory models, joining broad based, already successful digital structures—like the Humanities, Arts, Sciences, Technology Advanced Collaboratory—with deep, local focus on “on the ground” cultural

content, stakeholder investment, and institutional buy-in. In this way the, we might combine the strengths of groups like HASTAC (diachronic participation, accessibility, and swift communication), regional humanities centres (a focus on the history and culture of unique areas), and formalized and direct networks of researcher transfer and enrichment (postdocs, graduate student transfers, and sabbaticals that explicitly engage various communities in research). This project will outline these three areas as they currently exist in the Cascadian Bioregion, and, via primary consultations and theoretical research, suggest ways to productively move forward in building a multimodal regional collaboratory. I will end by extrapolating how the lessons of regional integration in this region might provide a model for regional collaboration in other contexts, especially the Australasian one.

The Cascadian Bioregion, extending roughly from Alaska south to Oregon and east to the Rockies, is an established center for the study of digital cultures, the teaching of digital methods, and cross-institutional collaborations. Given these investments, as well as the numerous existing commercial, governmental, and academic ties connecting the region, the Pacific Northwest is an ideal locale to consider what a multimodal collaboratory focused on regional issues might take. Such a collaboratory would differ markedly from larger efforts such as HASTAC, DH Commons, centerNet, or the Digital Research Infrastructure in the Arts and Humanities (DARIAH) in its focus on regional issues in the arts, humanities, and social sciences; in its ability to garner widespread support from those who value places and their histories because of long association; in its ability to leverage the existing academic, governmental, and cultural ties that form an integral part of any region's identity; in its commitment to preserving and facilitating access to local cultural materials; and in its ability to successfully enact public knowledge production and dissemination through close association with a community.

Preparation for this presentation at Digital Humanities Australasia 2014 will include discussions and recorded interviews with key academic and cultural stakeholders in the region, investigation of best practices for multi-modal collaboration, and consultation with experienced parties outside the region working in cultural memory and digital research institutions.

13:20- 14.50 **Sessions**

5.1: Digital History & Virtual Heritage II

Presentations

A 'Living Archive' for Circus Oz: investigating digital techniques to deploy cultural memory

David Carlin

RMIT University, Australia

This paper presents the methods and results of a major multi-year digital humanities research project, the Circus Oz Living Archive. The project has developed a new digital platform for an interactive video archive, with an interdisciplinary team of design, media, computer science and performance studies researchers working in partnership with the internationally significant Australian performing arts company, Circus Oz, as well as the Australia Council and the Performing Arts Collection of the Victorian Arts Centre Trust. The project was the subject of an invited keynote address at the UK's national Arts Marketing Association conference in 2013, and is the focus of a scholarly book, *Performing Digital*, to be published by Ashgate in 2014. Many performing arts companies around the world own highly valuable — but equally

vulnerable — informal archives of video and other materials documenting their performance history. However, until now, there has not been a digital platform available that could, firstly, enable agile access, interaction with and contextualisation of the unedited long-form video recordings that typically are the key form of performance documentation; and secondly, be designed so a performing arts company could itself build, maintain and deploy its own digital archive as a key part of its ongoing creative and organisational practice. The Circus Oz Living Archive project, in designing, developing and trialling a solution to this specific industry-driven need, has identified and addressed a complex and interconnected set of problems through an interdisciplinary set of ‘generative digital humanities’ (Burdick et al 2012) methods. In the Circus Oz Living Archive project, an iterative process of design and dialogue has included the writing of code, the development of an API, user testing and an ongoing series of workshops involving the project group and Circus Oz artists and staff, together constituting the knowledge-producing activities of the research laboratory. The results are expressed through the operations — the ‘performance’ — of the Living Archive site itself, which, having been released publicly in 2013 in beta, continues to be enriched as Circus Oz annotate and add metadata to over 400 performance videos tracing their 35-year history, as well as shoot and upload new performance videos.

This paper will articulate findings from the Circus Oz Living Archive project suggesting innovative strategies for industry and community-oriented digital humanities projects focused on cultural memory, as well as resistances and contradictions that have emerged around questions of openness, control and intellectual property. It will discuss what might be called the ‘humanities design’ approach developed through the project, as a forum for conceptual problem-finding and -solving, which incorporated the ‘coding as sketching’ techniques adopted by interaction designer and programmer Reuben Stanton. The questions raised by the project can be theorised as constituting the issues at stake in attempting to broaden a performance company’s mode of performance from the staging of physical action to include the scripting and manipulation of digital code. And at the same time, viewed from the opposite direction, in attempting to adapt and reconsider the concept of the ‘digital archive’ so that it reflects and strengthens the unique culture and practices of a contemporary circus.

'The First YouTube War': The U.S. wars in Iraq and Afghanistan as case studies in digital history

Michael Kerr Gisick

Australian National University, Australia

By 2007, and probably well before, even the smallest American outpost in Iraq had a room where soldiers with rifles slung across their backs hunched over dusty, malware-infested Dell computers to check in, via satellite, with the world wide web. What, I often wondered, were these young men and occasional women at the centre of future history reading? My newspaper stories, I hoped.

Alas, based on occasional snooping through browser history tabs, it appeared journalists like me could not compete with social media, even then. Far and away the most popular destination, outpacing even email, was Myspace, the social network that had recently passed Google as the most visited site in America. This was how soldiers in what has been called the ‘first YouTube war’ communicated with the outside world, creating in the process a vast, loosely connected, multi-media, semi-public, self-referential, often ironic, and sometimes heartbreaking archive of what it was like to be a soldier.[1]

The historical discipline has largely defined itself in relation to the archive. However, as historians interested in the digital sphere have recognized for some time, changes in communications technology have reconfigured the notion of the archive in fundamental ways, creating a “vast cabinet of curiosities” that is networked rather than hierarchical, relatively democratic rather than elite or institutional, and both easy to access and very hard to organize.[2] It is also, as the case of MySpace makes clear, simultaneously immediate and highly perishable.

This paper uses the wars in Iraq and Afghanistan as case studies for examining the practice of digital history. I agree at the outset with scholars such as Roy Rosenzweig, Tim Hitchcock and Paul Arthur regarding the urgent necessity of engaging in this practice. Given the high visibility and self-evident historical importance of the wars in Iraq and Afghanistan, they provide an ideal forum for exploring the ways digital culture has reshaped the historical archive, and the opportunities and challenges this presents for historians and other interested in its study.

I organize this paper along three fronts, using specific examples for each. The first relates to ‘official’ sources. How are the military’s own historians dealing with the transition to digital materials? How do the official documents likely to be available compare to past wars? What about ‘un-official official’ sources, specifically the trove of documents released by wikileaks? Secondly, I look at images, especially videos posted to sites like YouTube. A popular means for soldiers to communicate, these represent one way digital media can ‘democratize’ the archive, but are they useful as historical sources? What problems do they present in terms of identification and verification, and how can these be dealt with?

Finally, I look at written communication: email, blogs and social networks. How have these been regulated or censored? How perishable have they proved? How can researchers deal with privacy issues?

Framed by the problems facing contemporary military historians, this paper’s focus on key issues in the wider field of digital history and its use of real world examples should make it of interest to a broad audience.

[1] Noam Cohen, ‘Through Soldiers’ Eyes, ‘The First YouTube War’, *The New York Times*, May 23, 2010, accessed Sept. 13, 2013, http://www.nytimes.com/2010/05/24/business/media/24link.html?_r=0.

[2] Paul Longley Arthur, ‘Toward a Global Digital History,’ in *Global Media, Culture, and Identity: Theory, Cases and Approaches* ed. Rohit Chopra and Radhika Gajjala (New York: Routledge, 2011); 184.

The “Play It Again” Popular Memory Archive: a Game History and Preservation project

Darren Peacock², Helen Stuckey², Melanie Swalwell², Angela Ndaliansi¹, Denise de Vries²
1University of Melbourne, Australia; 2Flinders University, Australia

Games are one of the most significant cultural forms of our times and yet they are poorly documented in Australia and New Zealand. As conversations about the cultural significance of digital games become increasingly sophisticated, debates about what institutions will keep and preserve and what to focus collecting efforts on – the hardware and/or software, the artefact/object, or documentation – continue. One area that has not yet been fully explored is the potential for player memories to contribute to building games history, to become a resource for both institutions and current and future researchers. Knowledge about the history of games is overwhelmingly held by private collectors and fans, with ephemera and other primary sources located amongst the general public. This paper presents and discusses the Popular Memory Archive (PMA), an online collaborative portal of the “Play It Again” game

history and preservation project. As well as providing a way to disseminate some of the team's research, the PMA taps into what is, effectively, a collective public memory by providing a technique for collecting information, resources and reminiscences from the public about 1980s computer games.

The Popular Memory Archive seeks to balance a history of production, in the specific national contexts of 1980s Australia and New Zealand, with a history of use and reception. In researching the history of production, we have sought to compile information on as many locally-produced 1980s games titles as possible. The Play It Again project has identified more than 900 locally-written titles (700+ from Australia and 200+ from New Zealand), and the team has selected a shortlist of 50 or so titles for the PMA. This paper outlines the methods and approaches used to present these games in an online archive designed for user contribution and feedback. The Popular Memory Archive database will facilitate enable the capture and long-term preservation of a range of perspectives about the games and provide a comprehensive source of information about them. The paper will explain how the data has been categorised and tagged to cater for different themes to be discovered and appended.

A central concern that motivated the creation of the archive was driven by the belief that digital games are more than inert code; they come to life in the act of play. Collecting games and other artefacts and preserving them is thus only part of the construction of a history about games. The PMA is designed to work with online retro gamer communities and fans, and this paper also reflects on how the PMA is also an active research method for collecting the memories of those who lived and played their way through this period.

5.2: Infrastructure & Collaboration

Location: Seminar Room 1

Presentations

Integrating research in the humanities: the HuNI Virtual Laboratory

Deb Verhoeven², Toby Burrows¹, Anne Cregan³, Alex Hawker⁴, Kerry Kilner⁵

¹University of Western Australia, Australia; ²Deakin University, Australia; ³Intersect; ⁴VeRSI; ⁵University of Queensland

This paper will provide a comprehensive look at the HuNI Virtual Laboratory, presented by members of the project team and the Steering Committee. HuNI is a major new infrastructure service for humanities researchers in Australia, funded by NeCTAR and developed by a consortium of thirteen institutions led by Deakin University.

HuNI ingests and aggregates data from a total of 28 different Australian datasets which cover a wide range of disciplines in the humanities and creative arts. HuNI also provides a number of online research capabilities for humanities researchers to discover and work with the large-scale aggregation of data. These capabilities enable researchers to create, save and publish selections of data from HuNI; to analyse and manipulate the data; and to export the data for reuse in external environments.

The paper will cover the following topics:

- The international context: infrastructure services and programmes for the digital humanities in Europe and North America;
- The HuNI Virtual Laboratory project: the goals, aims and benefits of the NeCTAR project; the structure, management and resourcing of the project; sustainability;

- The HuNI data model: how it was developed, what it covers, and how it is used in the data aggregation process;
- Capabilities: how researchers can use the HuNI VL to discover and work with HuNI data;
- Users and stakeholders: how HuNI is involving humanities researchers, stakeholders and the wider community in the design and uptake of the Virtual Laboratory;
- Data provision: looking at HuNI from a data custodian's point of view – what needs to be done to contribute data to HuNI; how HuNI can augment and enhance the original data supplied by a dataset custodian.

The HuNI Virtual Laboratory directly addresses two of the major themes of the Digital Humanities Australasia conference: “Working with Data”, and “Building the DH Community and Presence”. HuNI is a major new development which engages with humanities and creative arts scholarship in the broadest sense and will be of significant interest and value to Australian humanities researchers.

The Human Communication Science Virtual Laboratory (HCS vLab)

Steve Cassidy², Dominique Estival¹, Peter Malcolm Sefton¹, Jared Berghold³, Burnham Denis¹
¹University of Western Sydney, Australia; ²Macquarie University, Australia; ³Intersect Australia

The Human Communication Science Virtual Laboratory (HCS vLab) is a UWS-led project, funded by The National eResearch Collaboration Tools and Resources project (NeCTAR), an Australian Government Super Science project, regrouping almost 50 active researchers from 16 institutions., 10 such Virtual Labs across Australia. The project builds upon 2 previous UWS-administered projects: the 2000-member Human Communication Science Network (HCSNet, ARC, RN0460284) and the 30-investigator, 12-institution Big Australian Speech Corpus (ARC LIEF, LE100100211) and the ANDS funded Australian National Corpus project led by Griffith University.

The main purpose of the HCS vLab is to provide an environment that will foster interdisciplinary research in Human Communication Science (HCS). While HCS is a broad field which encompasses speech science, speech technology, computer science, language technology, behavioural science, linguistics, music science, phonetics, phonology, sonics and acoustics, research is often conducted in isolation within each discipline. Too often the data sets used in research are difficult to share between researchers and even more between disciplines; tools are rarely shared across disciplines. HCS research in Australia, and the development of successful real-life applications, demands a new model of research, beyond that of the isolated desk/lab/university bound research environment. The HCS vLab environment aims to eliminate the waste involved in repeated unshared analyses, provide the impetus for new collaborations, encourage new tool-data combinations, and improve scientific replicability by moving data and tools as well as the analyses conducted with these into an easily accessible, shared environment.

Architecturally, the HCS vLab comprises a repository for heterogeneous data under a standardised metadata framework based on RDF, providing discovery services that allow researchers to create data sets that can be fed to a wide variety of research tools via a rich Application Programming Interface (API). Another major component is a workflow engine which allows data to be fed through a series of processing steps which can be stored and re-used. The HCS vLab will also orchestrate the creation of virtual environments including virtual

servers pre-loaded with both a set of tools and data, as well as virtual High Performance Computing clusters.

This presentation will first cover the architecture of the HCS vLab and give examples of its use across different kinds of data (text, audio, video) with a variety of tools. These include both 'point and click' pre-configured tools and a range of full programming environments in which data can be automatically marshalled for further processing. Examples include the Python-based Natural Language Toolkit (NLTK) for text processing and EMU-R on the R-stats platform for speech processing and analysis. We will present in more detail the variety of corpora that have been made accessible and discuss the tools that are available for analysing these data sets, emphasising the novel use of some of these. The presentation will then report on experiences with new kinds of interdisciplinary research and demonstrate some research scenarios.

We will also discuss the potential for this approach and architecture to be adopted more generally in the digital humanities world, showing how new data and tools can be imported into the Virtual Lab environment, and how the tools can be used on data anywhere.

Designing system integration across the Heurist, FAIMS and HuNI projects

Ian Johnson, Artem Osmakov
University of Sydney, Australia

XML has been a key technology for integration of Heurist (<http://HeuristScholar.org>) into the NeCTAR-funded FAIMS (Federated Archaeological Information Management System) and HuNI (Humanities Networked Infrastructure) projects. This paper will outline the integration workflows for these projects, including ontology mapping and the use of XML as a transport mechanism.

Heurist is a broadly applicable collaborative database tool which allows researchers to design and build their own richly linked databases through a web interface, without the need for specialist technical assistance. It can query and render its content in a variety of XML formats, as well as html reports, maps, timelines and simple descriptive statistics.

The FAIMS project takes advantage of cheap Android tablets to place digital data collection devices in the hands of every member of a field team. Rather than requiring a technical expert in the field to integrate data from multiple devices, the FAIMS system synchronises data automatically with a central database on a field server or laptop. The tablet app is highly configurable, using XML-based schemas to define both the data structures and tablet interface. The system has the potential to revolutionise fieldwork across many disciplines, from environmental fieldwork to archival research.

The writing of FAIMS schemas is, however, highly technical, limiting potential users of the system. We have therefore developed Heurist functions to write out the FAIMS XML schemas and related files directly from a Heurist database structure, allowing simple web-based creation of a FAIMS database and mobile app. A job requiring days of high level technical development has been reduced to a job the researcher can do on their own in a matter of hours. Heurist will also import a FAIMS database (creating both structure and data). It provides a number of XML output functions allowing data to be integrated into aggregation and sustainable repository systems. Specific export functions developed for FAIMS upload directly to The Digital Archaeological Record (<http://tdar.org>) repository system. Heurist also

exports file-based packages of XML data, and an Open Context (<http://opencontext.org/>) connector is under development. Heurist has several functions for building new databases from existing templates or by selective choice of structure from existing databases. We have developed a template compatible with the HuNI core ontology, which provides a mapping allowing the export of a set of XML files which are harvestable into the HuNI aggregate. This makes any new Heurist database searchable within the HuNI Virtual Laboratory, provided it is based on this template and authorised by HuNI. We have also developed mappings for existing databases, notably the Dictionary of Sydney (<http://dictionaryofsydney.org>). In this paper I will discuss the design decisions taken to develop integration across these systems. I will analyse the FAIMS workflow, from initial project design through forms creation, field data collection and synchronisation, to analysis and publication on the web and eventual deposit as a self-documenting archive with limited or no interoperability with other archives. I will contrast this workflow, with the integration of a more traditional Humanities database, the Dictionary of Sydney, into the HuNI aggregate. This workflow is characterised by the mapping of archival research and writing to a shared ontology, resulting in interoperability rather than sustainability.

5.3: BOF: All of the Records, All of the Time

Location: Seminar Room 2

Presentations

All of the records, all of the time

Morgan Glen Strong^{1,3}, Meg Travers²

¹University of Western Australia, Australia; ²State Records Office of Western Australia, Australia; ³Western Australian Museum, Australia

In 2012 the Western Australian Museum (WAM) and State Records Office (SRO) started parallel projects to implement open source Collection Management Systems at their institutions. Although each institution's system was built on a different software framework, one of the major outcomes desired from this project was a method for the information from each organisation to not just be combined, but deeply integrated. The institutions saw numerous commonalities between the data, not just thematically, but temporal and spatial synergies. For example, journals and maps from SRO could be combined with specimen data from the WAM to give deeper meaning to both sets of artifacts. The more the datasets were examined, further exciting relationships could be suggested by mining and combining these data. Utilising the Apache Solr search index and some sophisticated modules within Drupal, we are developing new methods to automatically integrate the datasets, and then provide visitors with a generous interface to help build their own connections and stories from our collections data. Through our websites, we will provide a window into the state's cultural and scientific heritage, but with a richer picture than either institution could provide on its own. This "Birds of a feather session" will facilitate a discussion to explore how cultural organisations can design their interfaces, and choose the tools to allow us to interlink our collection systems so that our visitors can deeply interrogate our rich datasets. The project is actively in development, allowing us to demonstrate lessons learned as well as the prototype of the finished product. It is our intention to release our developments back to the community and ensure code is available for others to reuse.

15:20pm - 16:50pm **Sessions**

6.1: Digital History & Virtual Heritage III

Location: University Club Auditorium

Presentations

Creating a Database of Religious Nonconformity and Performance in Britain (circa 1620-1680)

Alison Anne Searle, Ian Johnson
University of Sydney, Australia

This paper will present the model of entities (including agents, bibliographical and performative works, artefacts, events, groups and persons) and relationships that is central to my interrogation of religious nonconformity and dramatic performance in seventeenth-century Britain. This digital modelling exercise is applied to interdisciplinary data within the humanities, integrating religious, literary, dramatic and geographical aspects of early modern British culture that are often viewed as disparate or inherently antithetical. It offers a new and illuminating way of understanding religious nonconformity and performance, highlighting the heuristic potential of complex databases in humanities research.

My research analyses the ways in which religious nonconformity was practised in seventeenth-century Britain and its relationship to dramatic performance. These two core areas of cultural activity are often perceived to be in diametric opposition. I am developing a theoretical approach that analyses the inter-relationships between these activities, focusing on spectatorship (or a politics of audience), space, dissimulation, conversion and belief. These are being defined and interrogated through a series of historical case studies covering England, Ireland, Scotland and Wales.

Utilising the open-source software, Heurist (HeuristScholar.org), the database is currently being designed in collaboration with Dr Ian Johnson (University of Sydney), the system's designer, to model a unique network of relationships within the project's data. Heurist is not simply a mechanism for storing and managing data. It is crucial to the development of the project's research methodology, integrating key concepts, case studies and a plethora of individual items (including manuscripts and GPS-located images). The abstract modelling of connections between actors, events, places and items integral to the design of the database has generated questions central to the unfolding direction of my research and to the way in which I define key concepts. These might not have emerged, or would have been posed differently, if I were operating within a traditional humanities context without the digital framework.

One of the key project outputs will be a searchable online data repository of known spaces (theatres, village greens, pubs and so on) used for nonconformist religious activity and other forms of dramatic performance in seventeenth-century Britain. It will integrate locations and records of performances with transcriptions of archival material, visual images and short video-clips of performance spaces. The database is initially being developed as a research tool. At the end of the project it will be released as a freely accessible scholarly resource; the data will also be archived in a sustainable, internally documented standards-based format within a suitable institutional repository.

Mandelbrot, Braudel and Taleb: fractal geometry and randomness in history

Toby Burrows

University of Western Australia, Australia

The school of historians associated with the journal *Annales d'histoire économique et sociale* was the most influential twentieth-century movement in historical scholarship (Burguière 2009). It first emerged in the late 1920s, and by the 1950s had become primarily identified with the methodology known as “histoire serielle” – the gathering and analysis of long-run series of statistics on prices, wages and population – with Ernest Labrousse as its main proponent. This laid the foundation for the kind of quantitative history which became widespread in the 1960s and 1970s.

Another highly influential *Annales* historian of the same period was Fernand Braudel, famous for his large-scale histories of the Mediterranean and capitalism in the early modern world. While Braudel made effective use of statistics, he was wary of pushing quantitative methodologies too far in a determinist direction. Braudel favoured what he called “histoire totale”, and played a key role in the development of what is now called “historical anthropology”. In this, he was drawing on the ideas of the founders of the *Annales* School, Lucien Febvre and Marc Bloch, about “mentalities” – a broad and relatively vague concept involving the reconstruction of the mental universe of people in past societies.

In 1957, Braudel approached the mathematician Benoit Mandelbrot with an offer of a senior position and his own research institute at the prestigious *École Pratique des Hautes Études* in Paris (Mandelbrot 2012:194). Mandelbrot had been using historical financial data to experiment with new mathematical approaches to the analysis and modelling of random occurrences. As it happened, Mandelbrot rejected Braudel’s offer and moved to IBM in the United States, where he went on to discover fractal geometry and make a major contribution to chaos theory.

In more recent years, Mandelbrot’s work on financial data has been taken up by Nassim Nicholas Taleb in his provocative and popular books on chance, randomness, “black swans” and alternative histories (Taleb 2010). This paper will examine the relevance of Mandelbrot’s and Taleb’s methodologies to historical research, and particularly to standard forms of quantitative statistical analysis. It will also discuss the broader implications of these methodologies for a narrative discipline like history, and their possible applications in the emerging world of Linked Data.

Burguière, André, *The Annales School: an Intellectual History* (Ithaca: Cornell University Press, 2009)

Mandelbrot, Benoit B., *The Fractalist: Memoir of a Scientific Maverick* (New York: Pantheon, 2012)

Taleb, Nassim Nicholas, *The Black Swan: the Impact of the Highly Improbable* (London: Penguin, 2010)

The Western Australian New Music Archive: Digital distribution of a music culture

Lisa MacKinney¹, Cat Hope², Lelia Green³, Tos Mahoney⁴

¹State Library of Western Australia; ²Edith Cowan Univeristy, Australia; ³Edith Cowan Univeristy, Australia; ⁴Tura New Music

“The technical structure of the archiving archive also determines the structure of the archivable content even in its very coming into existence and in its relationship to the future.

The archivization produces as much as it records the event.” Jacques Derrida, 1995) In 2013, the Western Australian Academy of Performing Arts (WAAPA) at Edith Cowan University (ECU) and Perth organisation Tura New Music partnered with the Australian Research Council, the State Library of Western Australia, the National Library of Australia and ABC Classic FM to develop a digital archive of Western Australian new music activity from 1970 to the present day. The Western Australian New Music Archive (WANMA) seeks to discover, collect, collate, digitise, store and disseminate music recordings, video documentation, scores and other evidence surrounding Western Australian music of an experimental and exploratory nature. The definition of new music for the purposes of WANMA will be that formulated by the Australian Music Centre, which includes “notated composition, electroacoustic music, improvised music (including contemporary jazz), electronica, sound art, installation sound, and multimedia, web and film sound.” (AMC n.d.). The curation of WANMA will be guided by and confront the challenges provided by such a broad definition, enabling explorations of the rich inter-relationships between creators, performers, events, genres, venues and organisations in our musical heritage. The project is unique in way the archive shares as well as perpetuates content, through a program of concert performances of music from within the archive, but also of material for the purpose of growing the archive. Working and sharing data with project partners ABC Classic FM and the National Library of Australia, and linking material with the digitization of the Australian Music Centre’s collection also locates the Western Australian project within a broader, national context. This paper discusses the processes involved in the early stages of the implementation of the archive at the State Library of Western Australia. The challenges of finding an appropriate digital copyright arrangement, the style of database and its interface as well as the definitions used to collect and manage materials for the archive are discussed.

6.2: Geo-Spatial Methods

Location: Seminar Room 1

Presentations

Mapping the Routes and Roots of American Deportation

Ethan Blue¹, Zephyr Frank²

¹University of Western Australia, Australia; ²Stanford University

This paper combines mapping and visualization technologies with qualitative archival research and narrative presentation to plot the spatial dynamics evident in the United States’ system of mass deportation in the early twentieth century. It pays particular attention to the history of transcontinental ‘deportation special’ trains, and the experience of those who were forced aboard them. These reconfigured railroad cars gathered so-called ‘undesirable aliens’ from around the nation—disdained for their poverty, political radicalism, criminal conviction, or insanity: all compounded by maligned national and ethno-racial difference—and conveyed them to ports for exile overseas. First developed to ensure the efficient expulsion of post-Exclusion Chinese and non-citizens with mental disabilities, the trains were mobile, carceral spaces, and their history reveals the deportees’ journey as a process through which national territory, political sovereignty, and community—three defining features of modern nationhood—were created and contested.

In this paper we draw on U.S. Department of Immigration annual reports as well as archival case files for our qualitative and quantitative data. In addition to a historical analysis of the emergence of American mass deportation, we will describe the methods we have employed to visually represent large amounts of data—migratory routes of those who traveled to the United States, as well as what we call the spatial ‘networks of capture’ that ensnared them, and which fed deportees into the tributary and main lines of the deportation’s spatial flows. In so doing, we will make novel arguments about the history of global migration and American state formation, as well as methodological contribution to digital humanities and the application of visualizations and mapping technologies for interpreting historical data.

Geo-lexicography. how to improve digital humanities research by means of geo-visualisation infrastructures

Eveline Wandl-Vogt

Austrian Academy of Sciences, Austria

Geo-visualisation offers a lot of meaningful tools to improve research within natural and applied sciences as well as every days business. In this presentation the author will focus on the role of geo-visualisation within digital humanities research.

She will starting by introducing shortly the european german speaking dialectal lexicograhny and the state of the art. Her research will be exemplified by the Dictionary of Bavarian dialects in Austria (WBÖ): funding of the institute in 1911, collection until the 50es, publication since 60es), which was about to follow the digital turn: digitisation of base materials since the 90es, online base materials since 2010, online dictionary since 2012. She informs about the ongoing semantic turn (WBOE @ SKOS and @ LOD), both, digital and semantic turn under the management of the author.

This part of the presentation focuses the transformation from a traditional dialectal dictionary project into the digital humanities research by introducing into both, the main data and meta-data structures of these projects as well as european networks like CLARIN, DARIAH, COST and EUROPEANA to be connected with.

She introduces on the theme of geo-data in these kinds of digital humanities projects, due to the fact, that much information is inherently spatial in dialectology and lexicography.

The author will give an overview of the tools developed or used in the project framework of the Database / Dictionary of Bavarian dialects in Austria electronically mapped (dbo@ema) for several types of data and meta-data, going to be Austrian CLARIN/DARIAH-showcases, that focus on the main three issues of geovisualisation: visualization, anayzation, navigation.

Finally, she will point out how these tools aid the dictionary compilation, corpus studies and data access. Vice versa the geoinformaticans are driven by lexicographers needs. Concluding she will introduce into the model of geo-lexicography, opening a new view on the data and exploring spatially structured (dialectal) data.

The map as a source for memory in the Patagonia Austral

Gustavo Urbano Navarro
Unpa, Argentine Republic

Until the mid-twentieth century, the Patagonia Austral region had a poor national preeminence and the national as well as local media were silent about certain episodes, later retrieved by those who committed themselves to start generating a fair policy memory. The social shaping of the region was structured upon the logics of subordination inherited from territorial distribution and crossed, at the same time, by different migrating currents that contributed to the framework of its identity construction.

Both, memory “traces” of the individual identity that keep personal objects from a subjective selection, and memory “traces” of the social identity that protect objects assumed as patrimony, institute a way to “make history just from below”, that is to say, fostering a greater participation in its production, by focusing on the life of groups or people, sub-represented or ignored in the traditional historiography.

This paper intends to collaborate in the construction of Patagonia Austral memories, from considering documentary sources regarded as articulating axes of the reconstruction of public, daily life and their subjectivities, offering an “extra-centric perspective” to hegemonic interpretations.

Our proposal is to examine and view the changing vision around a group of crucial concepts, such as space, land, mapping, discovery, contact, native status, and other expressions, in order to comprehend the contrast between the speech of spatial understanding from the first aborigines in the Patagonia and that from the Europeans’. We think that the material for the authentic aboriginal voices is scarce and nearly all of them are recorded within European texts and documents. Is it the historical map and its modern real cartography about GIS elements to counteract this dichotomy?

A Methodology for Mapping Instagram Hashtags.

Tama Leaver, Tim Highfield
Curtin University, Australia

Social media platforms for content-sharing, information diffusion, and publishing thoughts and opinions have been the subject of a wide range of studies examining the formation of different publics, politics and media to health and crisis communication. For various reasons, some platforms are more widely-represented in research to date than others, particularly when examining large-scale activity captured through automated processes, or datasets reflecting the wider trend towards ‘big data’. Facebook, for instance, as a closed platform with different privacy settings available for its users, has not been subject to the same extensive quantitative and mixed-methods studies as other social media, such as Twitter. Indeed, Twitter serves as a leading example for the creation of methods for studying social media activity across myriad contexts: the strict character limit for tweets and the common functions of hashtags, replies, and retweets, as well as the more public nature of posting on Twitter, mean that the same processes can be used to track and analyse data collected through the Twitter API, despite covering very different subjects, languages, and contexts (see, for instance, Bruns, Burgess, Crawford, & Shaw, 2012; Moe & Larsson, 2013; Papacharissi & de Fatima Oliveira, 2012).

Building on the research carried out into Twitter, this paper outlines the development of a project which uses similar methods to study uses and activity on through the image-sharing platform Instagram. While the content of the two social media platforms is dissimilar – short textual comments versus images and video – there are significant architectural parallels which encourage the extension of analytical methods from one platform to another. The importance of tagging on Instagram, for instance, has conceptual and practical links to the hashtags employed on Twitter (and other social media platforms), with tags serving as markers for the main subjects, ideas, events, locations, or emotions featured in tweets and images alike. The Instagram API allows queries around user-specified tags, providing extensive information about relevant images and videos, similar to the results provided by the Twitter API for searches around particular hashtags or keywords. For Instagram, though, the information provided is more detailed than with Twitter, allowing the analysis of collected data to incorporate several different dimensions; for example, the information about the tagged images returned through the Instagram API will allow us to examine patterns of use around publishing activity (time of day, day of the week), types of content (image or video), filters used, and locations specified around these particular terms. More complex data also leads to more complex issues; for example, as Instagram photos can accrue comments over a long period, just capturing metadata for an image when it is first available may lack the full context information and scheduled revisiting of images may be necessary to capture the conversation and impact of an Instagram photo in terms of comments, likes and so forth.

This is an exploratory study, developing and introducing methods to track and analyse Instagram data; it builds upon the methods, tools, and scripts used by Bruns and Burgess (2010, 2011) in their large-scale analysis of Twitter datasets. These processes allow for the filtering of the collected data based on time and keywords, and for additional analytics around time intervals and overall user contributions. Such tools allow us to identify quantitative patterns within the captured, large-scale datasets, which are then supported by qualitative examinations of filtered datasets.

6.3: BOF: Digital Humanities Afield

Location: Seminar Room 2

Presentations

Digital humanities afield

Suzana Sukovic¹, Peter Read², Diana Hodge³

¹St.Vincent's College, Potts Point; ²University of Sydney; ³University of South Australia

Digital humanities are normally associated with organisations such as universities, museums, galleries and historical associations where we think of categories such as creators, students and users. In reality, the use and creation of digital humanities work happens in a continuum across sectors. Knowledge and skills required in digital humanities are being developed in a variety of contexts well before one becomes a digital humanist. Multipurposeful electronic editions, which provide exploratory environments for a variety of users and their needs, are a holy grail of digital humanities. Ideally, skilled users from all walks of life would create their understanding of the topic and use the material to create other works. However, a required skill set to participate in knowledge-building cycle online is becoming increasingly difficult to define as technology and digital practices blur traditional boundaries.

In this session, we wish to discuss users and creators of digital humanities outside the confines of academia and major cultural institutions. Who are these users/creators? What skills do they need? How is their work relevant to the mainstream digital humanities? We will use examples from secondary and tertiary education and from the community engagement with research data to discuss these questions.

The starting point for Suzana Sukovic is multimedia work by high school students which ticks all major boxes of concern to digital humanities – interpretation of text through digital media, blurring boundaries between creative and analytical, and ever expanding skill set. She proposes transliteracy as a framework for considering a range of related skills. Her model for integrating transliteracy in the curriculum, training and professional development programs, targets users and producers of digital humanities. It also provides a framework for embedding transliteracy skills relevant to digital humanities in a variety of learning contexts.

Diana Hodge considers issues of transliteracy in the tertiary environment. Librarians are familiar with the concept of information literacy and how these skills can be taught and how students can apply them across a range of platforms and formats. The term ‘transliteracy’ is still relatively new to academic librarians; librarians are however familiar with the idea that 21st century students need to be literate across an ever widening range of mediums. This ranges from archival documents, data, traditional text, audio, social media and art in physical and virtual formats. Can academic librarians help students from first year through to research degree level become transliterate? What specific skills, beyond the well-known standards for information literacy, should be included in a students’ repertoire for transliteracy?

Peter Read discusses uses of a historical community website. Who uses it? And How? historyofaboriginalsydney.edu.au is designed particularly for Aboriginal people as a way of returning individual and community histories back to the people who made them. It is also meant to be used by schools and colleges, and by Sydney people generally. In this discussion he will share some of the feedback the project team has received, and from whom, and how. The results are surprising. In this session, we wish to start discussions about digital humanities as a continuum. We believe that looking afield brings some relevant insights home.

18:30 **Optional book reading & signing with author Alexander McCall Smith**

UWA Octagon Theatre, UWA Crawley Campus

Main Conference - Friday, 21st March 2014

8:45 – 9:15 Registration

9:15 – 10:15 **Keynote - Dr Anthony F. Beavers**

Professor of Philosophy and Director of the Digital Humanities Lab, The University of Evansville, USA.

Tony Beavers is Professor of Philosophy and Director of the Digital Humanities Lab, The University of Evansville, USA. Tony works in the developing area of computational philosophy, an approach to the discipline that involves using computers to make philosophical discoveries that are not readily available with traditional argumentative methods and that also tests philosophical theories for computational tractability. Most recently, his efforts have involved showing that ethics may be more easily computable than might first appear and that this may well compromise our sense of values. Tony also works on network models that attempt to show how standard databases can be transformed into predictive mechanisms to help humanists make textual and historical discoveries.

Tony has lectured widely on these topics in the United States and Europe. He was also a Digital Humanities Fellow at Indiana University, 2008-2009, a fellow at the U.S. National Endowment for the Humanities Advanced Topics in the Digital Humanities at UCLA in 2010 and a co-planner for a second NEH Institute at the University of North Carolina, Charlotte in 2011. Anthony is the recipient of the 2012 World Technology Award in Ethics for his study of how machine ethics relates to our moral sensibilities. He recently served a term as President of the International Association for Computing and Philosophy.

Location: University Club Auditorium

10:45- 12.15 **Sessions**

7.1: Visualisation I

Location: University Club Auditorium

Presentations

I-Weave: interactive 3D costume visualisation

Rachel Fensham¹, John Collomosse²

¹University of Melbourne, Australia; ²University of Surrey, UK

iWeave was a collaborative research project in the field of digital archiving that involved the novel use of 3D data capture and interactive technology. It undertook the digitisation of fragile material objects, specifically original costumes from a dance archive that dated back to the 1920s. The initial research included reconstruction of original dances based on photographic footage before digital rendering from footage captured in a studio. The computer animated dances were then modified to support a Kinect interface, where users could try on the historical costumes virtually, as well as interact with the choreographic re-constructions.

At a practical level, the project also considered the corporeal histories of textiles, and dance practices, as the products of patterned repetition. It aimed to ask whether the material trace of the costume rendered digitally as interactive, and wearable in a new form and new media, could provide access to earlier modes of production, or different forms of perception about movement? In this sense, we explored how the digital archive might shift from recording past objects towards an interaction with the materiality of actions that retain a unique and specific movement dynamic.

While highlighting the possibilities for 3D capture and interactive technology to facilitate the interactive archiving and display of fragile historical items, the research also focused on the intimate relationship between costumes and embodiment. The dynamic properties of dancers who understood the tensile properties of clothing and these earlier expressive dance forms became vital to understanding the flow, shape and pattern of what constitutes digital kinesthesia, one of the most difficult qualities to understand in computer animations. 3D graphic representations and animation structures had to reproduce the surface properties of the garments via multi-layered and innovative algorithmic functions, and then to repattern coded data of movement to approximate the complex calculations made by a dancer in animating a costume. The Kinect software provided a limited range of variables for user interaction that further constrained the choreographic potential of the digital costume, however, the traces of the dancer and the subtle texture of the garment produced a new range of aesthetic visualisations.

The paper will show footage of the project design and its outcomes in visualising choreography and costumes; as well as consider the potential for 3-D movement to be transmitted digitally with benefits for user interaction.

Developing Collaborative Serious Game for Japanese Cultural Learning in 3D Metaverse

Mitsuyuki Inaba, Michiru Tamai, Ruck Thawonmas, Koichi Hosoi, Akinori Nakamura, Masayuki Uemura

Ritsumeikan University

This paper introduces our ongoing research and practice on collaborative serious game for learning of Japanese culture in 3D Metaverse. Serious game is a concept that utilizes game-based interaction and dialogues for the purpose of education and training. 3D Metaverse is a platform for constructing immersive virtual spaces with digitized objects and architectures on the Internet. A controllable virtual character in the space so called "avatar" enables the visitors to the 3D space to participate in embodied social interaction with other avatars. Therefore, 3D Metaverse as an infrastructure of serious game has a great potential to be a platform for inheriting and sharing traditional cultural heritage from the perspective of hands-on experience and situational learning.

First, we demonstrate our Metaverse environment in SecondLife(SL), which is the most popular Metaverse service. The space includes Japanese tangible cultural properties such as Shinto shrine, Buddhism temple, and Noh stage. It also has virtual museums for Kimono costume or Yuzen textile design. Various intangible cultural heritages, such as Noh performance, are also digitized and preserved in the space.

Second, we elucidate our experiments of collaborative serious game for learning Japanese culture on SL. In this game, pairs of Japanese and international students respond to quizzes

about the reason behind of Japanese cultural habits. Since these quizzes require in-depth knowledge about culture and values, it is difficult task even for Japanese students. Therefore, the participants had to refer to guidebooks, search web sites, and exchange their answers or interpretations on Japanese cultural habits.

Third, we describe the result of our analysis on the interaction and learning processes through the collaborative serious game in SL. In these learning processes, the students encountered their lack of knowledge about Japanese cultural practice and became to consider them in an equal position. As a whole, the collaborative playing between newcomers and old timers of Japanese culture makes the situational cultural learning between the participants a fun thing to do.

Finally, we argue about the result of the debriefing session by participants. Both foreign students and Japanese students answered that they could get better understanding of Japanese cultural practice through the serious game. Some participants also mentioned that the cross-cultural interaction during the game enhanced participants' awareness of not just the other's culture but their own culture. This result suggests that serious game could be a tool for promoting not only Japanese cultural learning but also cross-cultural understanding among players.

Simulating Ancient Cultures with Crowds of Virtual Agents

Tomas Trescak, Anton Bogdanovych, Paul Arthur, Harold Short
University of Western Sydney, Australia

Using 3D visualisation in reconstruction of lost sites of high historical significance has become a popular way of communicating the results of years of research conducted by archaeologists and historians to general public. Initially, such works were predominantly focused on reconstructing destroyed or partially destroyed architecture (e.g. Roman Colosseum). Such reconstructions help to simulate significant historical sites in all their former glory and facilitate appreciation of what remained from that glory (normally in the form of ruins). With modern advancement in research in development, we are now reaching the stage when reconstructing a heritage site can become relatively cheap. A procedural approach to generating historically informed designs of high complexity can be automated by design grammars, so that a large city can be created in a matter of days. One of the well known examples of using this approach in historical reconstructions is the Rome Reborn project, where a virtual reconstruction of the entire city of ancient Rome in the period of 320 AD was procedurally generated.

While 3D simulation of buildings and artefacts provides a unique possibility for general audiences to examine the architectural details of the heritage site it still does not help an observer to understand how this site has been enacted in the past. Without being able to see ordinary people performing their daily chores and rituals the observer is unable to immerse in the actual culture of the reconstructed society and have a complete picture about their way of life. It is possible to simulate such people using so-called "virtual agents". These virtual agents are essentially autonomous computer programs that are represented by human-like 3-dimensional figures (called avatars) that move around the reconstructed environment and simulate ancient citizens of the simulated site. Modern video games are a good illustration in regards to possibilities that arise with employment of such virtual agents in simulating human behaviour. But the cost of developing video games is enormous. For example, the estimated cost of developing Crisis 3, one of the popular modern video games, is \$66 Million. It's hard to

imagine such level of spending when it comes to historical simulations, so populating a historical environment with virtual agents needs to be automated.

Aiming to achieve cost saving, some researchers don't model their societies at the level of individual agents, but employ "virtual crowds". A number of crowd simulation and crowd generation approaches appear in the literature but hardly any of them advance beyond having avatars moving around and carrying objects with them.

In our project, we created a novel approach that can significantly decrease the cost and effort required for simulating everyday life of ancient inhabitants of virtual cities, while still capturing enough detail to be useful in historical simulations. Using our approach it is possible to design a small number of individual avatars and then automatically simulate a substantially large crowd of virtual agents, which will live their lives in the simulated city, perform chores and rituals as well as other routine activities that are consistent with their social status. The key novelty of our approach that enables simulating such sophisticated crowds is the combination of physiological needs - for generating agent goals, emotions and personality - for choosing how to fulfill each goal and genetically informed propagation of appearance and personality traits - to propagate aspects of appearance and behaviour from a small sample of manually designed individuals to large agent groups of a desired size. The usefulness of our approach is demonstrated by applying it to simulating everyday life in the ancient city of Uruk, 3000 B.C. See the video of our prototype at: <http://www.youtube.com/watch?v=NbEjOfjsz1w>

7.3: BOF: Debate - That Literary Studies Needs More Graphs, Maps, and Trees

Location: Seminar Room 2

Presentations

Debate: That Literary Studies Needs More Graphs, Maps, and Trees

Brett Hirsch¹, Philip Mead¹, Jason Ensor^{2,4}, David Large³, Katherine Bode⁴, Hugh Craig⁵, Roger Osborne⁶

¹University of Western Australia, Australia; ²Murdoch University, Australia; ³University of Sydney, Australia; ⁴Australian National University, Australia; ⁵University of Newcastle, Australia; ⁶University of Queensland, Australia

Is 'big data' the next big thing in literary studies, or are we losing sight of the texts that are -- or once were -- the focus of our scholarly endeavours?

This 'Birds of a Feather' session brings together experts from a variety of areas within the wider discipline of literary studies to debate quantitative approaches in the field and the politics and pragmatics associated with their application. The debate is intended to be a fun and informative way to engage with the fear and fascination surrounding the contentious application of quantitative methods in literary studies, a field in which traditional modes of qualitative analysis remain dominant, and to open up discussion about wider issues of interdisciplinarity, research assessment and value, the place of digital humanities in curriculum, etc.

The Australia-Asia debate format will be used, in which two teams of three speakers will debate the topic "That Literary Studies Needs More Graphs, Maps, and Trees" in the negative and the positive, each giving 5 minute speeches in turn before the first speaker of each team delivers a 'reply' speech in summation. A chairperson will keep time, and declare the winning

side as adjudicated by the audience. (For more details on this style of debating, see <http://tinyurl.com/kw3whgx>).

The formal debate will take up 40 minutes, opening up the remaining 20 minutes of the session for wider discussion amongst the speakers and audience.

13:15- 14.40 **Sessions**

8.1: Visualisation II

Location: University Club Auditorium

Presentations

‘It’s not how big your data is, it’s what you do with it that counts’:

Towards ‘expanded data’ in the Kinomatics Project

Deb Verhoeven¹, Colin Arrowsmith², Bronwyn Coate¹, Alwyn Davidson¹

¹Deakin University, Australia; ²RMIT University, Australia

The Kinomatics Project explores, analyses and visualises the industrial geometry of motion pictures (the study of ‘kinematics’ is often referred to as ‘the geometry of motion’) and is one of the first ‘big data’ studies of contemporary cultural diffusion. Its examination of global film flow rests on a large dataset of showtime information comprising more than 120 million records that describe every film screening in 48 countries over a 12-month period as well as additional aggregated box-office data. The multi-disciplinary research team behind Kinomatics includes specialists in film industry studies, GIS, geo-visualisation and cultural economics.

The project is based on the premise that films can be understood as cultural goods that are distributed both between ‘territories’ or markets and across the globe according to industrially unique spatial patterns and temporal flows. Seeing film in this way invites us to explore the industrial aspects of movement and location but it also invites reflection on our use of these large datasets.

For example, understanding the dynamics of global film exhibition and distribution demands an appreciation of scale in both the film industry and in a data-driven approach to its study. Exhibition and distribution of film titles is typically defined at the scale of an individual cinema, at a regional (‘territorial’) level, through national regulatory practices, and internationally. Our data is also able to be ‘scaled’ so we can aggregate information to give insights into the international impact of micro-attributes such as the role of individual stars in the success or failure of films or the impact of Awards ceremonies on global release patterns. But it is our ability to integrate this data with other forms of data – demographic data, financial data and climatic data for example, that offers the most exciting opportunities for new understandings of film industry logistics.

This presentation will argue that it is the adoption of an ‘expanded’ approach to cultural data, rather than ‘big data’ per se, which is especially valuable for cultural analytics projects such as Kinomatics. We will present spatio-temporal visualisations for representing heterogeneous data in order to push the scale and breadth of analysis of cultural datasets themselves.

Visualisation of immersive video as a tool to study embodied cognition

Khadeeja Ibrahim-Didi

Edith Cowan University, Australia

Our ability to access information is both conscious and unconscious and often impacted by the perceptual idiosyncracies that are unique to the bodies we inhabit (Lakoff & Johnson, 1999). The situated and immersed experiential knowledge associated with using one's body creates sensory awareness of 3-dimensional space and time, and although tacit, facilitates most cognitive activities and thought (Davis, Sumara & Luce-Kapler, 2000;; Maturana & Varela, 1987; Varela, Thompspon & Rosch, 1991). The role of bodily perception in regulating our interactions is recognition of the implicit engagement that occurs behind the overtly intended actions that occur within educational settings. Additionally bodily actions work with other communicated aspects to influence meaning (Suitner, 2012).

Immersive 360 degree video offers a unique, interactive, insider perspective to examine contexts that were previously unavailable for use in educational and training contexts. The video recording captures the setting in all directions. Further, it offers the capability to interact with the recorded video. Importantly, the video is captured from a single position that can emulate the act of looking around provides a bodily sense of 'being there'.

The way in which the video is viewed can also add to this feeling of embodied presence and develop the viewers' awareness of how they may feel if placed in a similar context. The captured video can be viewed in different ways at many levels of immersion. The first option is to view the video through a computer-based 360 manipulable video-player. Such players allow users to view and interact with the video on a computer screen, but with the ability to pan around in 360 degrees within the captured context at any one time and to zoom in, in any one direction. This view affords the perspective of embodiment that mirrors being able to look around, but does not provide a full-bodied sense of being there. Another option is to view the 360-video on a half-sphere iDome which fills the user's field of view, including the peripheral vision, removing the 'frame' of the rectangular screen and providing a sense of depth and of actually being 'within' the scene. The level of immersion experienced within iDome settings have been reported by some to create very significant, almost visceral, embodied responses to the video. The ability to manipulate the video to 'look' around a space where the field of vision is completely filled with the video scenario produces an authentic, embodied response.

Placing learners in such settings to reflectively examine scenarios from an embodied perspective allows them to examine their own responses to different educational settings and develop metacognitive control over their own responses, practice the regulation of their own responses in specified settings, and increase their ability to 'notice' aspects of practice that have particular significance – all within a safe environment. It affords opportunities to develop a sense of body awareness and the ability to 'read' the information that is produced in the moment (Schön, 1987), and to use that to complement the more intentional and cognitive aspects of reflection (Kinsella & Pitman, 2012) and inform their actions on the fly - much like what a teacher would have to do in a real classroom setting.

This presentation focuses on three projects that focus respectively on teacher professional learning, the development of teacher social emotional competence and evaluating cultural diversity training all highlighting how embodied educational research is facilitated through the visualisation capabilities of 360 degree video.

The problems and potentials of visualisation from a visual communication designer's perspective

Kate Sweetapple

University of Technology Sydney, Australia

Visualization tools – maps, charts, network diagrams and graphs – are being used in Digital Humanities to assist in the analysis and understanding of data sets, and to communicate findings. But not all digital humanists are convinced by the appropriateness of these adopted tools. Johanna Drucker argues that by borrowing tools from the natural and social sciences the digital humanities have unwittingly borrowed the conceptual and visual language of scientific positivism, which does not suit the “fundamental epistemological values” of the humanities. (2012) She writes that these visualisation tools “carry with them assumptions of knowledge as observer-independent and certain, rather than observer co-dependent and interpretive.” (2011: 1) Drucker argues that human experience is not well served by existing digital tools of visualisation and calls for a radical reworking of the current conventions of graphical expression.

This paper argues that the field of visual communication design is well placed to assist in this radical reworking. Designers understand the rhetorical conventions used to display information and the ways in which to manipulate these visual strategies. Importantly, they are also keenly aware of the impossibility of neutrally represented information. At the core of the visual communication discipline is an understanding of the contingent relationship between form and content – of the fallacy that information can be independent of its means of expression.

To make this point the paper will be divided into two sections. First, a brief discussion of three information visualisation practices: scientific, journalistic and artistic. (Hall 2011) I will discuss how traditionally, visualisation was used solely as a scientific tool for discovery, a method for understanding and analyzing large data sets, and how more recently it has been used in journalism to simplify and explain data sets. [1] The third category is artistic practice. This category expands the aesthetic dimension of the visualisation and attends to affective aspects of communication, with the primary concern being the mode of representation. It is from this last category, I will argue, that a ‘radical reworking’ of information visualization may arise. As Peter Hall states, the role of artistic visualisation is to “offer new, alternative modes of representation,” and “call into question the claims of transparency, certainty, and objectivity embedded in the Cartesian language of the genre.” (2011: 184) What is being ‘called into question’ through this practice is the language of scientific positivism.

In the second part of the paper I will show examples of work that incorporates aspects of scientific and journalistic visualization but draws most heavily from ‘artistic practice’. This will include my own work around experimental cartography and that of contemporary designers, Jonathan Puckey and Stefanie Posovec. I will discuss how these emerging representational paradigms, and the creative and critical concerns of a visual communication designer, could assist in the creation of a more ‘humanistic lens’ through which to represent visualisations from the digital humanities.

Drucker, J. (2011) “Humanities Approaches to Graphical Display,” *Digital Humanities Quarterly*, Volume 5 Number 1, The Alliance of Digital Humanities Organisations <http://digitalhumanities.org/dhq/vol/5/1/000091/000091.html>

Drucker, J. (2012) “Representation and the digital environment: Essential challenges for humanists,” <http://www.uminnpressblog.com/2012/05/representation-and-digital-environment.html> (viewed 12 Sept 2013)

Hall, P (2011) “Bubbles, Lines and String: How Information Visualisation Shapes Society,” *Graphic Design: now in production*, Minneapolis: Walker Art Center, p.170 - 185

8.2: Literature & Literacies

Location: Seminar Room 1

Presentations

Literature in the database: archiving, standards and sustainability.

Maria Angel, Anna Gibbs, Peter Sefton

University of Western Sydney, Australia

Adelta is the Australian Directory of Electronic Literature and Text-based Art. The directory emerges from the Creative Nation ARC discovery funded project: Writers and Writing in the New Media Arts led by Assoc/Prof Anna Gibbs and Dr Maria Angel from the Writing and Society Research Group in the School of Humanities and Communication Arts at the University of Western Sydney. Its technical development is led by Dr Peter Sefton, eResearch Manager at UWS. It aims to produce an authoritative, relatively stable, descriptive and interactive archive of Australian digital media writers and writing. We use results from our research on emergent forms and genre taxonomies to provide up-to-date information that reflects developments in the field and current usage while building an historical record of work since 1990.

It has become clear that technical and conceptual development of the Adelta Directory are inseparable. New modes of archiving and access adequate both to new forms and objects and to new ways of working need to be developed according to standards that are both local and international in order to be sustainable over time in emerging national and international research infrastructures and changing disciplinary formations.

A key aim of the project is to establish interoperability between various partners in CELL (International Consortium for Electronic Literature) who are developing comparable databases. This consortium addresses the development of collaboration for the purposes of researching, publishing and archiving electronic works. Founding partners of CELL are the University of Western Sydney (Adelta), ELMCIP (University of Bergen, Norway), ELD (US); Po.Ex (Portugal), NT2 (Canada), University of Siegen (Likumed). Adelta and the CELL group use modern metadata best practice, as negotiated with CELL, aiming to be "Linked Open Data Ready". The presentation will report on innovative approaches to metadata standardisation across the international CELL consortium.

Initial technical implementation of this project takes the form of a registry of works of Australian electronic literature and text-based art, using an Open Source repository platform, Islandora. The choice of a repository platform means that the service can be expanded from a registry to a preservation and dissemination repository in the future, and can also be expanded into a research database covering other kinds of entities, such as people and critical works.

We also discuss sustainability for the project via formal governance arrangements with the UWS Library and e-Research committee, and a formal data management plan for the project that will ensure hosting and archiving to an agreed schedule. This approach is part of the university's investment in the digital humanities, building capability and generic infrastructure such as repositories to encourage sustainable research in the area.

From Users to Producers: the Theory and Practice of Fostering Undergraduate Digital Literacy

Shannon Rose Smith

Bader International Study Centre, Queen's University (Canada), United Kingdom

In his recent consideration of the way in which the digitisation of source material has transformed the study of the nineteenth-century press, *The Nineteenth-Century Press and the Digital Age* (2012), Victorian print culture scholar Jim Mussell highlights the need for instructors to engage students of these materials in discussions of not only their cultural and historical context, but also of the way in which this material is mediated to the student through a digital interface. Mussell argues that in order to fully engage critically with such material, both instructors and students need to move beyond viewing such digital objects as 'mimetic' or surrogate; in the process of digitising such material, it is 'transformed into something completely different' and thus should be subject to a critical evaluation that considers not only issues such as editorial decisions, but also the social and cultural context for the production of the digital object and the way in which access to it is proscribed (160). With such digital objects increasingly being incorporated into undergraduate study of humanities disciplines such as history, literature, and art history, developing strategies for fostering in undergraduate students the capacity to not only use such materials, but also to critically evaluate them is imperative.

How do we equip students to become active, critical users of the digital resources at the heart of many humanities disciplines? This paper will examine the design and implementation of a programme of undergraduate courses at the Bader International Study Centre, Queen's University (Canada) focussed on developing a heightened level of digital literacy among users of digital resources and the broader digital culture of which those resources are a significant part. It will evaluate some of the decisions that were made in the process and reflect on the course framework's central theoretical concepts of 'user' and 'producer'. This paper will posit a model for digitally literate undergraduate engagement with digital culture that draws not only on discussions of digital literacy by Richard Burniske (2008), Jim Mussell (2012), and Rodney Jones and Christopher Hafner (2013), but also on recent considerations of the participatory art movement -- perhaps best represented by Hans-Ulrich Obrist's ongoing installation *Do It* (1993-2013) -- by scholars such as Claire Bishop (2012).

Multimedia Essays as Tools For Cultural Mediation

Miguel Escobar Varela

National University of Singapore, Singapore

In this paper I define a multimedia essay as a digital artifact with the following qualities:

- 1) It was conceived to be navigated, searched and commented on with digital platforms in mind.
- 2) It is interactive and can be accessed in numerous ways. For example, through hyperlinked visualizations that allow for non-linear reading.
- 3) It contains non-textual elements such as video, images and sound.
- 4) It contains Natural Language Generation (NLG) algorithms that can re-accommodate the contents of the essays in response to search queries.
- 5) It includes interactive tooltip displays that present contextual information such as

definitions of glossary terms, translations, and bibliographical items.

6) Comments can be added to any part of the artifact.

My main interest is to investigate the ways these multimedia essays allow for users to access materials from cultural contexts with which they might not be familiar. In order to describe these possibilities in greater detail, I will refer to a project currently under development, a web-based PhD dissertation on wayang kulit shadow puppetry in Indonesia.

For my research, I have recorded and translated several videos of wayang kontemporer performances. These performances are re-elaborations of Wayang Kulit, Java's oldest and most prestigious tradition, that address topical concerns such as environmental threats, youth culture, politics and changing ideas about the role of art. In order to appreciate the significance of the aesthetic and thematic transgressions that make these performances so compelling, an observer needs to be familiar with the prevailing conventions in the cultural context where they were created. In order to contextualize the performances for people not familiar with their cultural dimensions, I am currently developing two websites:

- 1) A freely accessible online archive of two dozen full-length video recordings, complete with subtitles and contextual information.
- 2) A web-based dissertation where fragments of the videos are integrated into multimedia essays that provide interpretations of the performances based on ethnographic field-notes.

In this presentation, I will consider the implication of the expressive possibilities of interactive, multimedia interfaces for the interpretation of artworks from different cultural contexts. I will offer a perspective on the general implication of these features for Digital Humanities, with a special emphasis on Performance Studies and Anthropological perspectives on art.

HTTP 404 Not Found: Hoverboards, AI, or the Literary Utopia of the Digital Age

Paul Michael Lee Belanger
Deakin University, Australia

The dream of electronic literature has been around since we first started bumping bits and crunching bytes. It has gone through fits and starts, produced a few gems, and established something of a canon in the work of its early practitioners (even if inclusion is often more by dint of their having gone first than anything else). Despite this long history, there remains an overall dearth of contemporary electronic literature that rises above the fray of fan-fiction, teenage angst, theoretical self-indulgence, or the occasional electronic oddity. Eastgate's signature hypertext writing platform Storyspace isn't yet updated for 64-bit systems, some titles require legacy hardware and software to run, other platforms have similarly fallen by the wayside or never made it out of their testing stages, and many of our earliest web-based artefacts now return nothing but 404's and non-functional pages. Those platforms and artefacts that do work – old and new – often come off as either simplistic or overly complex in design, exhibit execution and/or writing that doesn't quite inspire, and may possess less than intuitive interfaces. Collaborative writing sites and communities, with some few exceptions, end up ghost towns or high schools. Resources end up missing, obsolete, or broken, and on it goes, leaving today's digitally inclined writers to toil in the dark corners of the net wondering what's become of their promised land?

This short paper looks at the current state of electronic literature, asking where its game-changer is hiding and if perhaps the medium is just now ready to start realizing some of the utopian goals its early proponents left us. It examines the difficulties and hurdles electronic authors (and creatives generally) face in the digital arena. Among these are problems of obsolescence, the ease of publishing coupled with a lack of 'reputable' publishing venues, steep learning curves, still existing economic and technological barriers, entrenched notions of literature, some aversion to collaboration, a series of increasingly difficult aesthetic questions, an absence from the classroom, and coming full circle, the existence of few examples and standards for writers to aspire to, toy with, and challenge. However, much as Fredric Jameson claims that utopias succeed in failure, might the sometimes spectacular failures of electronic literature have already succeeded in laying the foundations for a new type of creative writer?

Indeed, increasing digital literacy, a budding understanding and appreciation of new types of literature, the re-entry of some early practitioners into the field, the proliferation of book and book-like apps for touch screen interfaces, growing communities of coders and writers, the release of new writing platforms, and conferences such as this one provide several bright lights on the digital horizon. Drawing from my experience writing and developing *Into the Machine* (an ideologically driven hypertext narrative), the findings from my upcoming *Encounters in Hypertext* project at the AAWP's 2013 conference, interviews with educators in the field, and my experience navigating the obstacles of the medium, this paper offers up some modest proposals for advancing the practice of electronic writing, building electronic writing communities and publishing outlets, and helping newcomers find their way through a sometimes challenging terrain. In conclusion, it offers a guardedly optimistic outlook on the emergence of not just one game-changer, our long awaited Shakespeare, but of a whole generation of game-changers reaffirming utopia, digital or otherwise, as process not end.

8.3: BOF: Beyond Digitisation - Preparing Non-Digital Resources for the Digital World and Curation of

Location: Seminar Room 2

Presentations

Beyond Digitisation: preparing non-digital resources for the digital world and curation of digital resources in the GLAM sector.

Lise Summers², Lise Summers¹

¹Curtin University, Australia; ²State Records Office of Western Australia

Practitioners from the GLAM sector are faced with many challenges in preparing their material for use in digital humanities research. Much of their collections are in non-digital formats, and the majority of digital information relates to metadata about the collection rather than delving into data sets within the collections. Digitisation is an expensive and time consuming task, and digitising without the appropriate metadata or in the wrong formats may make all that effort unusable. This BoF will look at the skill sets and understandings required for GLAM professionals to support digital humanities research both now and in the future (or what to do if you are not Tim Sherratt). The discussions will draw on real world experiences in preparing archival collections for GovHack Perth and in teaching information studies at Curtin University, and will ask practitioners where to from here.

9.1: Visualisation II

Location: University Club Auditorium

Presentations

The Invention of Microscopes: Data Visualisations and Spatial Resolutions in the Humanities

Thomas Koentges

University of Victoria Wellington, New Zealand

Realising the limitations of the naked eye, four centuries ago researchers developed microscopes and telescopes which became influential research instruments. Since the seventeenth century humanity has improved these tools even further and we are now able to research objects that are only 0.1 nanometer small (the thickness of a human hair is around one-hundred thousand nanometer) or more than thirteen billion lightyears away (the Big Bang is dated to about 13.8 billion years ago).

These achievements alone are very impressive for a species that still fails to answer the very fundamental question of any research: that is, 'Why?' Although this question has been repeatedly asked for over three thousand years, the reason for the repeated failure to answer the question is the limitations of human resolving power: the human mind, is a powerful but highly limited research instrument. This is why researchers attempt to reduce complex problems to a model they can understand better or they take samples and then try to deduce predictions and assumptions. Although these tools offer some help and relief, they are imprecise, because the approach they take is a reduction of the complexity of a problem to something simpler, something we can understand. Given the intrinsic imprecision and limited nature of these tools, they cannot be the optimal approach to very complex problems.

Given that in recent times humanities is facing an explosion of research data, the question of how to research complex and abundant data is more important than ever. If we could invent an optical tool that helps to provide an overview of the whole complexity of the problem and if we learn to read the images and results of this tool properly, perhaps we can enhance human wisdom and understanding.

Katy Börner, inspired by the work and theories of Joël de Rosnay, provided a theoretical foundation for a new approach when writing her article 'Plug-and-Play Microscopes': 'Decision making in science, industry, and politics, as well as in daily life, requires that we make sense of data sets representing the structure and dynamics of complex systems. Analysis, navigation, and management of these continuously evolving data sets require a new kind of data-analysis and visualization tool we call a macroscope [...].'

Börner hopes that such a new tool will help explore the 'infinitely complex'. Science and even more so the humanities, are becoming increasingly interdisciplinary and data driven, and given the complex and fuzzy state of linguistic or historic data for example, there is certainly a need to apply such a macroscopic tool. In this paper I will present a small survey or selection of worldwide projects that research and visualise very complex data in the humanities, while I focus in particular on Social Network Analysis in a historical context. In conclusion I shall suggest some guidelines for constructing or working with a macroscope.

(Digital) Bread and Circuses: Reframing Ancient Spectacle for Different Screens

Anna Foka

HUMLab, The University of Umea, Sweden, Sweden

It is commonplace that screen-based communication – i.e. TV, cinema, computer screens and ubiquitous devices is continuously mediating cultures (Galloway 2004, Giaccardi et al. 2012). Digital reconstruction is the process of graphically representing ideas and objects (Wileman: 1993). This process, however, requires a conceptual picture to be transferred to in a graphical medium. This paper focuses on the potentials of a conceptual digital construction of a Roman Amphitheatre for multiple screens. I argue that while current ‘historically accurate’ digital depictions of Roman amphitheatres are limited to lifeless and sanitized aerial 3D models, a more innovative, multisensory and participatory reconstruction of entertainment sites for multiple screens can elucidate our understanding of historically and geographically remote social and cultural concepts.

I propose new methodological tools for generating discourses that add layers of understanding to our contemporary knowledge of the Roman spectacle. A participatory (embodied- tangible computing) and multisensory (sound and vision) digital recreation of a Roman amphitheatre (along the lines of Betts: 2009, Drucker: 2009, and Favro: 2006) can engineer deeper and constructive analyses of the dynamics and systemic operations regarding [ancient and current] popular entertainment. It can generate questions about the cultural and emotional context of ancient spectacle as well as the potentials and limitations set by our current technological grasp. It can further be applicable in research and education in order to anchor both ‘traditional’ research questions, as well as the importance of multiplicity within institutional material infrastructure.

9.2: Digital Editions & Editing

Location: Seminar Room 1

Presentations

Stable, sharable, referable, reusable digital scholarly editions (DSEs): future or fairytale?

Desmond Schmidt

University of Queensland, Australia

Scholarly editions have been around at least since the Alexandrian variorum editions and commentaries of the first century AD. And the modern scholarly edition has not changed so much in function or form from that early model. But how can such editions continue to be built in a digital world? Have the problems inherent in digital presentation and in the modelling of scholarly interaction with the text been solved to the same degree as in the print scholarly edition? Can such editions be created using general tools at low cost, that are stable and can be referred to, that earn money to sustain them, that can be easily repurposed, read, shared and published for the long term in a digital library? The uncomfortable truth is, none of these desirable things are currently possible.

Commercial formats like the eBook, scanned and OCRred copies of old books on Google, digital copies of open texts on Project Gutenberg can be downloaded or viewed online. Digital humanists have been outdone by the easy availability of such forms of digital edition, while

their own efforts can still be characterised mostly as experimental. So why not simply allow industry and commerce to supply the technology to create the DSE? The problem is that the forms of data with which scholars work - the text of the edition itself - and the kinds of interaction needed cannot be easily recreated using industrial tools. Some of the work at least must be done by digital humanists themselves because they are the only people who truly understand what needs to be done.

But in working out exactly what an interoperable and sharable DSE might look like in the future one approach may simply be to eliminate possibilities. Firstly, a sharable and stable DSE cannot be composed of computer code. This seems to be impossible because sharing would require the establishment of an international standard for a functional interface to something that is in constant flux. Secondly, a sharable DSE must supply coherent, readable texts without embedded markup codes, at least in one of its forms. The possibilities for interpretative choice of different ways to encode the same information, and the selection of which information to record are so great that it is now generally recognised such texts can never be standardised. Thirdly, a DSE cannot be a collection of views because the future possibilities for visualising texts are practically endless, and depend on software. All that is left for a stable DSE to be is thus a disposition of raw data, a file structure, that facilitates the basic functions of the scholarly edition: comparison, commentary, cataloguing and referring, markup and structure. However, technology can never be completely banished from such a simple model: even the concepts of file, directory and digital text are technologies in themselves. One approach, copied from the software industry, is to provide multiple equivalent formats for the same data so that multiple standard programs can read it.

Stability can be supplied by agreements with established publishers, universities and libraries to store data for the long term, and to provide editions with fixed URIs, such as handles that permanently refer to digital objects. But to make all this possible the DSE must be identifiable as a specific digital 'thing': detached, stable and able to be referred to, not hidden in the bowels of a customised website or commercial edition. Seeing the potential for such a form of data by demonstrating its possibilities may be the best incentive to developing standard practice through evolution rather than revolution.

The AustESE Workbench: Supporting Collaboration in Electronic Scholarly Editing

Anna Gerber, Roger Osborne, Jane Hunter
The University of Queensland, Australia

1. Introduction

The Australian Electronic Scholarly Editing project (2012-2013) is a NeCTAR-funded collaboration between the University of Queensland, University of NSW, Curtin University, University of Sydney, Queensland University of Technology, Loyola University, Chicago and the University of Saskatchewan. The project aims to develop interoperable Web-based services to support the production of electronic scholarly editions by distributed collaborators.

The AustESE project team has developed an Open Source scholarly editing Workbench that coordinates scholarly editing workflow and provides access to online tools to support editing tasks. The workbench encapsulates:

- a content repository that supports the creation, import, discovery, search, retrieval and re-use of bibliographic metadata, transcriptions and images that comprise an electronic edition;
- tools for editing and aligning transcriptions and images;
- collation tools that automatically detect and highlight variants between different versions of a work;
- annotation tools that enable scholars to create, reply to, search and browse scholarly annotations attached to transcriptions, variants and images;
- visualisation tools for exploring the content and relationships between transcriptions, variants, annotations and bibliographic metadata including Works, Versions, Artefacts, Agents and Events;
- publishing tools for generating online electronic editions and for exporting to standard publication formats.

The Workbench currently hosts several large-scale editorial projects, and supports the scholarly editing workflow for each project by managing roles, permissions, and project progress, as well as capturing provenance information.

The AustESE project will complete development in December 2013. In this paper, we report on the project outcomes by outlining and reflecting on the challenges that have arisen during the course of developing the technical architecture, underlying data model, and implementation, as well as describing editing projects currently using the Workbench. Finally we discuss how the foundational data model and architecture established by AustESE might be extended to support future editions.

2. AustESE Workbench

We have adopted a Service-Oriented Architecture, and have developed modular, reusable, and potentially distributed components that can be assembled and substituted according to the requirements of each scholarly edition project. We extended existing scholarly editing tools (e.g. Calliope) to enable their integration within our Drupal-based Workbench, and also implemented new open source software to bridge the gaps between existing tools.

3. Data Model

We have developed an OWL Ontology for electronic scholarly editing (ESE) that includes key classes: Work, Version, Artefact, Agent and Event. The ESE Ontology provides the data model used to organise the metadata and conceptual entities within the repository. The Open Annotation data model (Sanderson, Ciccarese & Van de Sompel, 2012) is employed for modelling annotations. We have extended the OA core model with specialised Annotation Motivations and Selectors.

4. Example Editorial Projects

The AustESE Project has maintained a close connection with Australia's scholarly editing community through members of the steering committee and through workshops conducted in Sydney and Brisbane. A number of editorial projects are currently managed through the Workbench, including Paul Eggert's Charles Harpur Critical Archive, Roger Osborne's electronic edition of Joseph Furphy's *Such is Life* and a critical edition of Joseph Conrad's *Nostromo*, as well as several other projects that have developed out of AustESE workshops.

5. Future Developments

The data that accumulates within AustESE-based collaborative electronic scholarly editions is represented using data models designed to support interoperability. This facilitates sharing, discovery and re-use of electronic editions and also provides potential for innovative methods of aggregated analysis and visualisation. Future work might include the integration of additional analysis and visualisation tools and the ability to render editions on innovative publication platforms.

9.3: Digital Evidence

Location: Seminar Room 2

Presentations

Whither Trautonium?

Meg Travers, Cat Hope

Edith Cowan University, Australia

The history of electronic music whilst comparatively short, is longer than most people (contemporary electronic musicians included) are aware. From the development of the Telharmonium in 1897 (considered to be the first electronic musical instrument), this new area of music quickly fielded a number of bespoke instruments developed for performers or composers to explore areas of sound previously unavailable to them using standard orchestral instruments.

The classical tradition of art music relies heavily on ongoing performances of works written many years prior, by new young performers as well as known exponents, but what of the works of composers of electronic music where the instruments and sound sources used may have fallen into disrepair, obsolescence, vanished, or modern technology changed the sound so much as to make it unrecognizable from the original intended instrumentation?

Beyond the traditional collections of manuscripts and recordings, the practicalities of re-performance of 20th century electronic musical compositions have not been widely considered, and there exists no tested methodology for assessing and archiving 20th and 21st century electronic musical instruments for use in re-performance of these works, though as greater importance is placed on this with the passage of time, the issue will become more apparent, and also more difficult to retrospectively resolve. To perform a curated collecting role for the performing arts, collecting institutions need to understand what needs to be collected, as well as developing methodologies for preserving if not the instruments, then the essence of their sound and performance techniques in modern technological equivalents.

One such instrument is the Trautonium, first developed in Germany in 1929. The instrument's principal performer and main developer, Oskar Sala, passed away in 2002 leaving behind many recordings of pieces composed for the instrument by himself and other noted composers of the time, and the instrument itself in a museum in Germany. However to all intents, the ability for people to perform these works on the instrument they were written for has been closed off. To understand what is required to successfully re-perform electronic music written for such an obsolete instrument, an attempt needs to be made to re-create the instrument and test it in the performance of works written for it when it was contemporary, assessing the sounds produced by the new instrument, its construction and playability using digital archiving methodologies to determine ways to ensure the long term continuance of the instrument.

Having undertaken explorations of the archiving of 1970's commercial synthesizers, this paper is about the continuation of that work in my Masters degree, taking a technical and practical look at the digital curation requirements of the instrument, how to create, and the fit-for-purpose, of a 21st century Trautonium.

Not just CSI: the use of digital forensics in research data curation

Leo Konstantelos, Anna Shadbolt, Gavan McCarthy
University of Melbourne, Australia

A room in half light, a scientist leaning over a supercomputer, a law enforcement officer, a despicable cybercrime to solve: the perfect setting for digital forensics. Or not quite so: in recent years, the use of digital forensics has emerged within the digital curation and archiving space as a methodology for capturing, identifying, analysing and recovering data from digital media so as to create and maintain evidential records of digital activity. Digital forensics extend the capabilities for curators and archivists to audit, establish custodianship and provenance, assess accountability and retain the authenticity of digital data in a forensically sound manner. Academic institutions offer a propitious environment to study theoretical and practical models for employing digital forensics in the management, curation and preservation of digital data that are the product of research activity. With continuing advances in digital humanities, the area of applicability of these models has shifted beyond "big (scientific) data": humanities research data are increasingly more complex, encompassing diverse digital objects and intricate interrelationships between sociotechnical structures. This paper presents the efforts of the University Library at the University of Melbourne to set up a digital forensics lab for curating humanities research data. We discuss the motivation and rationale for establishing the lab, alongside current and prospective objectives. In support of the application of digital forensics in humanities research, we demonstrate our methodological approach, process and outcomes from the forensic curation of the Australian Sound Design Project (ASDP) data. This paper concludes with insights on the use of digital forensics within digital humanities research as a means to demonstrably protect its evidential and cultural value now and into the future.

Stories in Stone: an annotated history and guide to the collections and papers of Ernest Westlake (1855-1922)

Rebe Taylor, Michael Jones
The University of Melbourne, Australia

Stories in Stone sets a new benchmark for web-based archive guides and historical web resources. It digitally preserves, annotates and disseminates the entire papers relating to three large stone implement collections formed by English amateur scientist Ernest Westlake from about 1870 to 1920 in England, France and Tasmania held in the Pitt Rivers Museum, Oxford and the Oxford University Museum of Natural History. Ernest Westlake had formal training in geology, but his wide-reaching research interests included psychical phenomena, cultural evolution and anthropology. Stories in Stone represents much of Westlake's life's work.

Digital archive guides are often confined to a single collection, and do not always capture that archive in its entirety. Even when item level archival description is included, historical contextual material is usually only present at collection and provenance level and is often sparse. *Stories in Stone* demonstrates the capacity for cross-disciplinary expertise: it includes five distinct archival collections – four digitised in their entirety – held in two institutions in order to collate in their entirety the papers relating to Ernest Westlake’s scientific research. In addition to standards-based, archival metadata, there is item level historical contextual explanation: how and why were the collections and papers created, how do they relate to other items documented in the guide, and beyond; to other archival collections, publications, events and people? *Stories in Stone* extends the archivist’s ‘traditional’ aim to capture the stories of the records in order to meet the historian’s primary aim: to reveal the stories in the records. Preservation meets enquiry to allow for independent searches of over 8000 images that are informed by more than 40,000 words of scholarly history writing. The digitised records are presented in an image viewer with the capacity to create customised citations of each image from EAD XML, including their location as paper records and their reproduction within the scholarly resource. *Stories in Stone* is not merely a means to view original records; it is the result of several years of primary historical research combined with the expertise of leading archivists specialised in cultural informatics.

One expert reviewer considers that *Stories in Stone* suggests a new direction in publishing digital collections: ‘that in the future this process (advanced with such tools as social networking and tagging) will increase. Digital collections are evolving from static lists of data to living venues for artifact and document based discussion ...I see the addition of external academic narrative as being very beneficial to researchers...’ (peer review for Rebe Taylor, ‘A Journey of 13,033 Stones: The Westlake Collection and Papers’, *Collections – a journal for museums and archives professionals*, Vol 8, No 1, 2012, pp 7-38).

In this paper, Dr Rebe Taylor and Michael Jones outline how they, with Gavan McCarthy, co-created *Stories in Stone*. This includes Gavan’s initial inventory listing in the Heritage Documentation Management System (HDMS) (a standards-based system developed by the ESRC over more than ten years to manage archival and artefactual documentation projects and to generate HTML and EAD finding aids for publication on the web) to the challenges of expanding the parameters of the HDMS to include Rebe’s extensive historical interpretation, as well as meet the citation requirements of the Pitt Rivers Museum. Rebe and Michael will also outline their plans for the future: to publish a monograph on Westlake and the broader history of anthropology and archaeology in Tasmania, utilising TEI and EAC-CPF to create a network of relationships to contextual entities and the published archive, linking historical notes to the records, and their related metadata.

15:50-16:00 **Closing Remarks – Professor Paul Arthur, aaDH President**

Location: University Club Auditorium