



THE UNIVERSITY OF
MELBOURNE

Highlighting the role of collaboration for digital HASS

Reviewing the Digital Studio Graduate
Internship Program

—
Professor Rachel Fensham – Assistant Dean, Digital Studio

Tom Vasey – Project Officer, Digital Studio

IDCC 2019 Lightning Talk





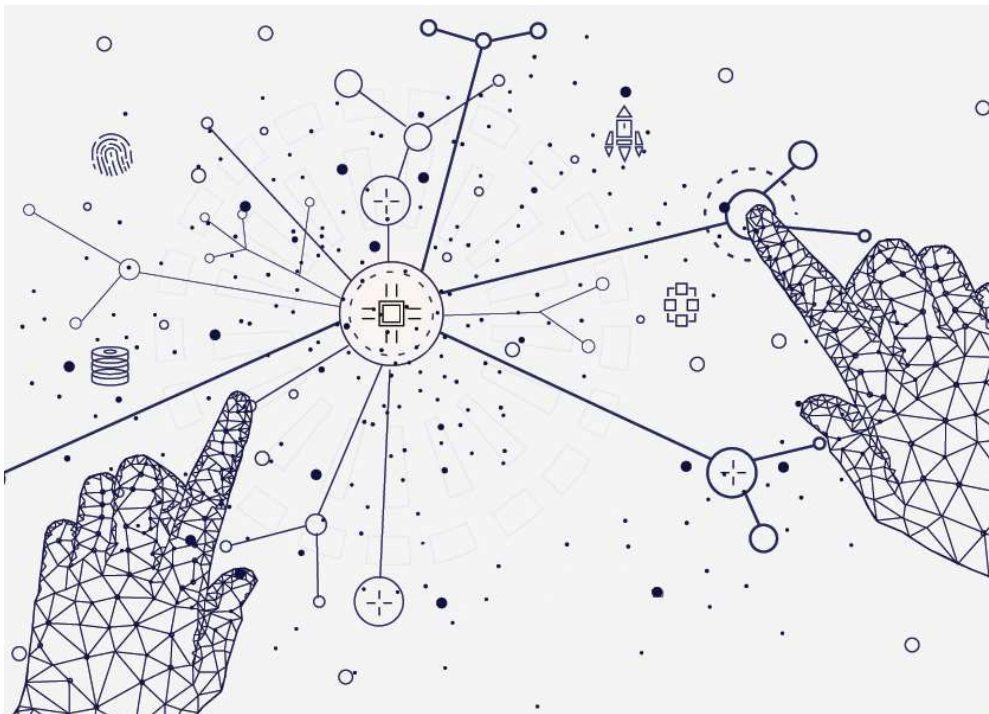
Digital Studio

The Digital Studio provides services, expertise, equipment and collaborative spaces for Faculty of Arts researchers and industry partners working in the digital humanities, arts and social sciences (HASS).

- Leading debates and discussion for digital scholars, whether critics, creatives or champions
- Community building and mentoring of the new generation of digital humanists
- Exhibition, curation and promotion of digital research in a mediated environment
- Piloting and enhancing HASS data projects and methodologies
- Fostering industry partnerships and digital innovation
- Membership and advocacy on University and national committees to advance digital scholarship and infrastructure
- Supporting international knowledge exchange, by hosting research fellows and partnerships



Graduate Internship Scheme



The Internships bring together graduate students, postdocs and academics from across the Faculty of Arts to work together on digital research projects.

- Supporting academics and research projects
- Mentoring a community of new digital humanists
- Trialling digital concepts and methodologies
- Providing insight into existing data practices and capacity
- A need analysis for future training and support



Structure of the program

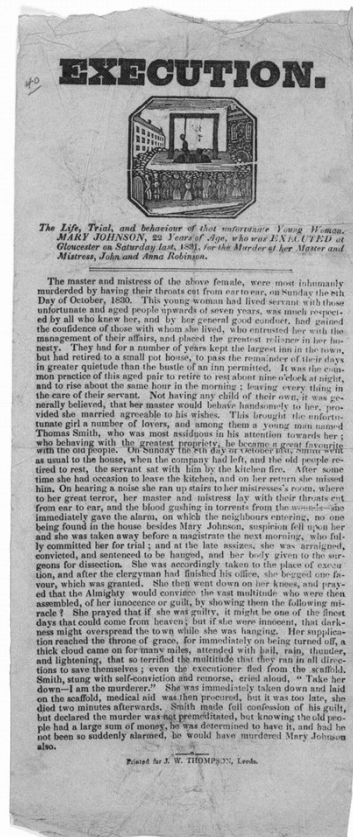
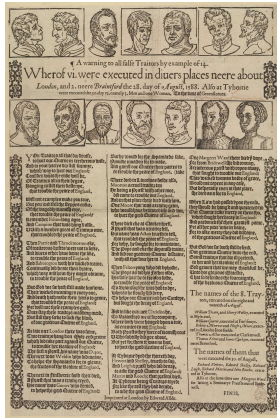
- 10 internships valued at \$2000 each (approximately 45 hours work).
- Each intern was assigned to an academic/researcher with a unique project, that can be enhanced by the expansion, exploration, or cleaning of its data.
- Training provided in basic coding and management of data in the academic research environment.
- Additional mentoring by the Digital Studio continued throughout the internship, interns were also encouraged to attend Research Platforms training.
- Interns were supported throughout the internship by the Digital Studio, with the support of SCIP, and encouraged to work in the Studio's collaborative digital research environment.
- A series of meet-ups were also held to help support collaboration and knowledge sharing among the intern cohort.
- Interns, supported by academics, were required to demonstrate in what ways the use of data from the project has been enhanced or expanded as part of the Internship Showcase.



Projects

Project	Digital skills	Digital tools
<i>Upgrade the Classics & Archaeology Virtual Museum</i>	Data wrangling; metadata coding; 3D modelling	Photogrammetry; Excel
<i>Execution Ballads of Pre-Modern Europe</i>	Data wrangling; metadata coding; visualisation	Omeka
<i>Network Analysis the Cross-Cultural Interaction of International Students</i>	Non-API data scraping; data wrangling; language analysis; visualisation	Polyglot.detect; Python
<i>War Words and the Evolution of Australian English</i>	Data wrangling; metadata coding; visualisation	Django Python Web framework; NLTK; Python (vis.js, dash, & pandas libraries)
<i>Captured and Captioned: Representing Family Lives on Instagram</i>	Non-API data scraping; metadata coding	Instagram metadata extracting tool (developed by SCIP)
<i>VCA Film and Television Digital Archive Project</i>	Metadata coding; audio visual metadata schema, ontologies, and standards	MediaFlux
<i>Melbourne Urban Directories: Putting People in Place</i>	Data wrangling; metadata coding; OCR correction	OCR data capture tool (developed by SCIP)
<i>Genealogies of Warruwi Community: integrating current materials</i>	Data migration; metadata coding; data wrangling; data analysis	Go (programming language); KinOath (genealogy application); XML; GEDCOM
<i>Digital Collection Engagement</i>	Methodologies and practices for cultural collection discovery and display; metadata coding	Cultural collection discovery platforms; online collection applications
<i>Visualisation for Database on Fatal Shootings of Drugs Suspects</i>	Data wrangling; verification; visualisation	Google sheets; Tableau

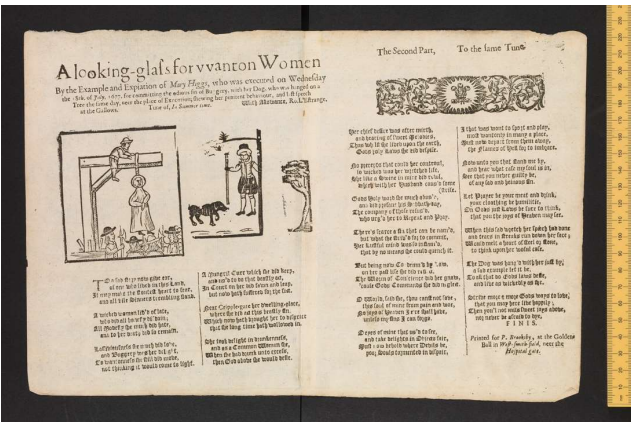
Project: Execution Ballads of Pre-Modern Europe



This project developed an Omeka database of execution ballads from pre-modern Europe dating from c. 1550-1900. A wide range of data formats are included in the site, such as images of the pamphlets and broadsides, transcriptions of lyrics, audio recordings, related artworks, and tune information.

Approach

- Sort, match and clean data
- Plan Omeka site structure
- Customise Omeka elements
- Batch import CSV metadata files
- Upload additional files (audio files, images)
- Create relations
- Refine website



Project: War Words and the Evolution of Australian English

War words dashboard

Enter search terms separated by a comma (regular expressions permitted)

Enter terms

Normalised values

Raw values



Minimum score



This project developed a website to facilitate the studying of *Aussie* magazine, a trench journal that was first published during 1918 in France. The website includes an interactive dashboard facilitating analysis of word frequencies and uses machine learning to find words and ideas related to a search term.

Approach

- Using fuzzy string matching algorithms to identify possible OCR errors
- Examine and visualise not just individual words, but entire concepts
- Using machine learning models to identify words with semantic or contextual link to the search terms



Common data curation and preservation challenges

The projects fell into three broad areas:

1. *Investigating and trialling metadata schemes and metadata coding processes and practices*
2. *Developing approaches for visualisation*
3. *Exploring language analysis; from content collection to processing*

Across the projects there were a number of common challenges:

- The importance of data processing workflows; where to begin and how to first approach raw data
- Value and limitations of existing metadata schemes versus project specific classification
- Appropriate and accessible language for documentation; i.e. how you describe and categorise OCR errors
- Planning for the longevity of platforms; interoperability and enabling future migration



Measuring success

Academic Feedback

86% of academics rated their involvement in the program as 'very positive'

80% reviewed the technical support and training provided as 'very good'

"I have nothing to criticise; I would just like to see it expanded!"

"great way to learn more about the digital studio and build a community of people interested in digital humanities"

"more 'peer learning' or a midway presentation amongst [interns]"

Student Feedback

71% of interns rated their overall experience as 'very positive'

43% of interns reviewed the technical support and training provided as 'good' or 'very good'.

"[Could explore] structured insight into the data pipeline, the process of going from raw research materials to a complete and useful digital asset"

"more attention should be given to how tools fit together within a broader digital research landscape"

"little attention was given to how we might structure our data on a level more fundamental than Dublin Core"



Lessons learnt

Advancing Data Curation in the University involves:

- **A generational shift** – differences between academic and graduate students experiences reflect divergent perspectives on the potential for expansion and development within digital humanities
- **Project start-up mentoring** - teaching how to plan and develop a data workflow, in order to address the difficulties of delivering training that is broadly applicable.
- **Integrated technology expansion** - once you create a community and people have the opportunity to engage with data curation projects and methodologies you need to be ready to develop more advanced training and capacity.
- **Localised access to relevant HASS support** – the critical path for navigating the limits of specific platforms and tools in an incomplete national infrastructure landscape.



THE UNIVERSITY OF
MELBOURNE

Thank you

For more details:

arts.unimelb.edu.au/research/digital-studio

[@digitalstudioUM](#)

Professor Rachel Fensham – Assistant Dean, Digital Studio

Tom Vasey – Project Officer, Digital Studio

IDCC 2019 Lightning Talk