



PANEL
SCHOLARSHIP OF THE FUTURE
DIGITAL SCHOLARSHIP FUTURES

Digital technologies are transforming how research is conducted. This paper will focus on one part of that transformation: in the humanities broadly, and in literary and historical studies in particular. Using a range of projects to illustrate each, I will discuss three broad trajectories in this context: regarding the new forms of access, curation, and engagement that digital technologies make possible. I will also show how each of these trajectories relates to my own research with the National Library of Australia's Trove's database of digitised historical newspapers. My focus will be on the possibilities that digital technologies hold out for humanities research, as well as the challenges this presents to the way we think about scholarship, and its requirements, aims, and outcomes.

Access

It has become almost commonplace to note that digital technologies dramatically transform our relationship to the humanities archive, in particular by increasing the range of materials we can access, and the ways in which are able to access them. Where once a scholar would have had to travel to the site where a document was located, perhaps make an appointment to see it (maybe even requiring a letter of introduction to do so), now many documents are available online. One example of such a resource is the British Library's 'Discovering Literature' portal, which provides high-quality digitised versions of rare editions of famous Romantic and Victorian literature, as well as a range of materials helpful in assessing and understanding these documents, such as historical essays and podcasted interviews with academics.¹

One consequence of the digitisation of our cultural record is a transformation in the type of evidence used in humanities research: scholars are increasingly seeking to explore that record not by looking at individual items, however far away they are, but by abstracting those documents and investigating them on a large scale. A notable project of this type is ‘Mapping the Republic of Letters’, based at Stanford University.² Drawing on digitised copies of letters to and from key Enlightenment figures, this project visualises correspondence networks of the period, and provides new historical insights on that basis. One such analysis includes a comparison of John Locke’s and Voltaire’s correspondence networks. We are accustomed to thinking of their correspondence networks as cosmopolitan; and Enlightenment writers positioned themselves in this context. In contrast, the visualisation suggests that Locke’s correspondents were heavily focused in the Anglo-Dutch area, while England was a surprising ‘coldspot’ for Voltaire.³

This theme of increased access to the documentary record via digital resources and methods is central to my own project. In the past three years, I have led a project to discover serialised fiction in nineteenth-century Australian newspapers. Literary and periodical historians have always recognised these newspapers as major publishers of literature in the colonies. But until now, the archive has been too extensive to test that assumption, except by indexing particular newspapers. Trove’s digitisation of a large number of historical Australian newspaper pages (almost 19.5 million at the time of writing)⁴ changes that situation profoundly. Its search interface allows us to use terms such as ‘chapter’ and ‘our novelist’ to identify fiction, while its Application Programming Interface (API) enables us extract the results of this process. So far I have discovered over 15,000 fictional titles published in Australia in the nineteenth century, including almost 10,000 serialised stories. Much of this fiction comes from overseas: from Britain and America, as might be expected, as well as Canada, France, Germany, New Zealand, South Africa, Russia, and beyond. But this project has also discovered a number of Australian titles not previously identified, and even brand new Australian authors: some of whom—such as Tasmanian author, John Silvester Nottage—are well

worth including in our existing literary canon. Without the type of access that digital technologies provide it would have been impossible to identify and amass this extensive corpus of fiction.

My discussion, thus far, has focused on documents—or textual artefacts—partly because that is the area I work in, but also because the technologies of access currently available are typically designed for use with text. However, the new modes of access that digital technologies enable extend beyond text to sound and image. In terms of visual collections, the Rijks Museum in Holland offers an innovative interface for exploration, providing access to astonishingly beautiful and high-quality digital images of all its objects.⁵ Researchers associated with the multi-institutional ‘Digital Music Lab’ are experimenting with new ways of analysing and visualising music, including exploring the internal dynamics of different genres.⁶ However, these new possibilities for access bring challenges for humanities research. Some especially pressing challenges include:

- What scale of evidence do we now consider necessary for a humanities argument?
- How do we interpret new types of evidence—such as data—in the humanities, and what should be done about enhancing the technical abilities of humanities researchers?
- How do we deal with the challenges presented by the potential gap between the access we think we are achieving and the access we actually achieve?

In respect to this last question, the ‘googlisation’ of society has accustomed us to believing that entering a term in a search box provides access to all information about that topic. In fact, complex algorithms manage and order that access, making it difficult to gauge the relationship between what we seek to access and the access we actually achieve.

Curation

Part of the answer to the challenges posed by new forms of digital access is the return of the curatorial role in humanities scholarship. In the nineteenth and early twentieth centuries, curation was a recognised sphere of activity

for humanities scholars. It was not unusual for a professor to be responsible for a collection of documents or artefacts in their area of expertise. The mid-twentieth century saw a decline in this practice as the number of academics expanded, as curation became a specialised profession, and as many fields of research moved from what might be seen as a scholarly paradigm to a research or theory framework.⁷ In the field in which I work, of literary studies, there has been a dramatic decline in people working on bibliographical and editorial projects in the latter part of the twentieth century.⁸ The entry of digital technologies into humanities research is bringing curation to the forefront of humanities scholarship once again. Multiple digital curatorial and editorial projects have arisen, although there is still little sense of what to call them. As Kenneth M. Price noted in 2009, in a claim that remains true today: ‘*Project* is amorphous; *archive* and *edition* are heavy with associations carried over from print culture; *database* is both too limiting and too misleading in its connotations; and *digital thematic research* collection lacks a memorable ring and pithiness.’⁹

Leaving terminology aside, many of the longest-running digital projects in the humanities are curatorial. One example is Price’s own Walt Whitman Archive, which brings together the published and unpublished works of this author, as well as multiple essays, biographical resources, and other commentary.¹⁰ Multiple other curatorial projects exist, and in fact, this process of bringing together related documents or artefacts—or rather, representations of those documents or artefacts—is one of the key ways in which humanities scholars are engaging with digital scholarship.¹¹ I am performing this curatorial role with my study of serial fiction in nineteenth-century Australian newspapers. It is one thing to amass a collection of over 15,000 fictional works, but the question remains: beyond analysing it for my own purposes, what does one do with such a collection?

While it would be a waste if this rich literary corpus were not made available, entering into the role of curator brings its own, major challenges, including:

- Who should perform this role—who has responsibility for humanities data?

- Should all data be freely available? What about culturally sensitive data? How do we manage the protocols of access in this respect, and who should take responsibility for enacting these?
- If data are to be made freely available, who takes responsibility for the sustainability of those datasets into the future?
- Should humanities scholars try, if possible, to give their data to some organisation or institution to manage, and if so, how can they be expected to deal with the complexity of that data and the challenges of making it available to others?

Where scientific datasets are often very large, humanities datasets are frequently small, multiple, and with hierarchically complex organisations. For instance, my serial fiction data describes the multiple different versions in which the same literary works are published in nineteenth-century newspapers. Humanities datasets thereby present a particular challenge for existing data repositories, which often—understandably—have a one-size-fits-all system for storage and discovery. Protocols are emerging for management of humanities data, but no clear consensus exists yet; accordingly, this emerging curatorial role for humanities scholars is yet to be controlled and sustained by agreed upon principles.

Engagement

Curation invites a form of engagement, in that scholars make a collection available to researchers and the general public. But digital scholarship also presents exciting opportunities to engage the community in the research process in new ways through crowdsourcing. In the scholarly context, crowdsourcing typically involves inviting members of the general public to undertake tasks that computers cannot yet do effectively (such as the pattern recognition involved in reading handwriting—or nineteenth-century newspaper print). Thus far, the sciences have been more successful than the humanities in this area, not least of all because they often have the funding required to enable such engagement. Zooniverse is a website for predominantly scientific crowdsourcing projects, that invites members of the public to—for instance, ‘help us identify individual humpback whales by cluing our computer algorithms in to patterns on their tails’.¹²

However, a number of humanities projects are taking advantage of crowdsourcing to enhance the scholarly environment and encourage public engagement with such research. Trove is recognised internationally as a major leader in the crowdsourcing arena, and it is estimated that text correction by members of the community has saved the Australian government more than 12 million dollars.¹³ Another humanities project that has effectively involved the community is ‘Your Paintings Tagger.’ Recognising that a large proportion of the oil paintings held by cultural institutions in Britain were in storage, the researchers behind this project digitised these paintings and invited the public to identify—or ‘tag’—features in them.¹⁴ In the future, these descriptions will be used to facilitate access to and analysis of oil paintings. The New York Public Library’s ‘What’s on the Menu?’ project is likewise using the power of the ‘crowd’ to transcribe the over 45,000 historical menus held by that cultural institution.

For my own project, much more fiction remains to be discovered: in already-digitised newspapers, and in the many hundreds of newspapers yet to be digitised. Not only had I hoped to implement my own crowdsourcing project to enable this ongoing discovery, but I had assumed it would be fairly easy to implement because Trove already provides some of its essential ingredients including: an excellent text-correction interface; an API for linking that interface to other projects; and a large, and engaged public, some of whom are already involved in finding and correcting serial fiction.¹⁵ For the moment, however, the challenges presented by this new form of engagement have prevented my move into this area. In particular, developing a successful crowdsourcing project requires significant expertise and funding. The tasks involved in such a project must be carefully scaffolded and integrated, and this is a difficult exercise requiring considerable experience and knowledge. Users are sophisticated in their digital engagement, and expect sophisticated interfaces to work with. These interfaces must be adapted to mobile devices as well as desktops, and be both attractive and easy-to-use to secure the interest of people for whom many other projects are available. Because of the expertise involved, crowdsourcing projects are expensive to develop, even if—once up and running—they can significantly extend the scope of research.

While the challenges of digital scholarship are considerable, so are its possibilities. In this respect, all of the emerging trajectories I have described—of access, curation, and engagement—require new skills, knowledge, and protocols. But they also present exciting opportunities, to enrich and invigorate the humanities, both for scholars and the general public.

-
- 1 British Library, *Discovering Literature: Romantics and Victorians*, <http://www.bl.uk/romantics-and-victorians>.
 - 2 Stanford University, *Mapping the Republic of Letters*, <http://republicofletters.stanford.edu/>.
 - 3 Stanford University, *Mapping the Republic of Letters, Voltaire and the Enlightenment*, <http://republicofletters.stanford.edu/casestudies/voltaire.html>.
 - 4 National Library of Australia, *About Digitised Newspapers and More*, <http://trove.nla.gov.au/ndp/del/about>.
 - 5 Rijks Museum, *Explore the Collection*, <https://www.rijksmuseum.nl/en/explore-the-collection>.
 - 6 Digital Music Lab, *About*, <http://dml.city.ac.uk/about>.
 - 7 For discussion of this historical curatorial role and its contemporary resurgence in the digital humanities see Anne Burdick et al, *Digital Humanities*, MIT Press, Cambridge, Mass., 2012, esp. pp. 17–18, 33–35, https://mitpress.mit.edu/sites/default/files/9780262018470_Open_Access_Edition.pdf, Open Access Version.
 - 8 See Carol Hetherington, ‘Old Tricks for New Dogs: Resurrecting Bibliography and Literary History’, in *Resourceful Reading: The New Empiricism, eResearch and Australian literary Culture*, eds Katherine Bode and Robert Dixon, Sydney University Press, Sydney, 2009, pp. 70–83.
 - 9 Kenneth M. Price, ‘Edition, Project, Database, Archive, Thematic Research Collection: What’s in a Name?’ *Digital Humanities Quarterly*, vol. 3, no. 3, 2009, <http://www.digitalhumanities.org/dhq/vol/3/3/000053/000053.html>.
 - 10 *The Walt Whitman Archive*, eds Ed Folsom and Kenneth M. Price. <http://www.whitmanarchive.org/>.
 - 11 Other curatorial projects in the field of literary studies include: Lehigh University, *The Vault at Pfaffs: An Archive of Art and Literature by the Bohemians of Antebellum New York*, <https://pfaffs.web.lehigh.edu/>; and King’s College London, *Jonathan Swift Archive*, <http://jonathanswiftarchive.org.uk/index.html>.
 - 12 Zooniverse, *Whales as Individuals*, <http://daily.zooniverse.org/2015/07/03/new-project-whales-as-individuals/>.
 - 13 Marie-Louise Ayres, *Digging Deep in Trove: Success, Challenge and Uncertainty*, 4 September 2012, <https://www.nla.gov.au/our-publications/staff-papers/digging-deep->

in-trove-success-challenge-and-uncertainty.

¹⁴ Public Catalogue Foundation, *Your Paintings Tagger*, <http://tagger.thepcf.org.uk>.

¹⁵ New York Public Library, *What's On the Menu?* <http://menus.nypl.org/>.