

Digital revolutions: The limits and affordances of online collections

There seems little need in a book like this to explain what Trove is. Anyone with an interest in Australian history will have lost countless hours amidst its 200 million digitised newspaper articles. Similarly, it seems obvious that Trove has changed the practice of history in Australia, but how? This chapter explores changes in the way Australian historical resources have been delivered online, and the implications of this for the way we *do* history.

As the COVID lockdowns in 2020 emphasised, the online availability of primary source materials makes historical research possible even when access to the originals is limited. But the undoubted convenience of being able to browse 200 years worth of newspapers at home masks other issues that historians have been slow to acknowledge and address. How do we discover relevant resources? What gets digitised and why? How can researchers use these collections to ask new types of questions? After more than 25 years, the web has its own history. What problems and possibilities has it brought to the way we understand the past?

Manufacturing discovery

For 30 years a group of academics and volunteers laboured to improve access to Victoria's past. Subject indexes to the *Argus*, Melbourne's 'major newspaper of record', were an important starting point for researchers, but there was a gap in the indexes between 1860 and 1909.¹ The Argus Index Project aimed to fill that gap. Struggling with limited resources, it took 15 years for the project to deliver its first set of indexes, covering 1860 to 1869.² Australian Research Council (ARC) funding from 2001 to 2005 accelerated the work, and the project promised the creation of a complete digital index of the *Argus* that would 'for the first time, make the contents of a late nineteenth century Australian newspaper widely available to researchers of Australia'.³ By 2006, the index for 1870 to 1879 was online, hosted by the National Library of Australia. Six years later it was gone.

In August 2009, the first version of the NLA's digitised newspapers site, later to become part of Trove, was made available online. It included the *Argus* from 1846 to 1945 (later extended to 1957) – digitised and fully searchable. It is easy to assume that Trove made the Argus Index obsolete. Why bother working your way through a detailed hierarchy of subject terms when you could simply type keywords in a search box? But the reality is more complex, and reflects a continuing tension in the way we find historical resources online.

The early years of the web were a time of lists. Before the rise of web-spanning search behemoths, experts and enthusiasts created subject portals – annotated lists of links to relevant resources. Just like the compilation of print indexes, directories, and bibliographies by generations of scholars and librarians, these lists helped researchers navigate a sea of information. However, the rapid growth of the web overwhelmed the list makers. Soon broad subject-based portals gave way to specialist databases, focused on

particular research topics. In 1994, the Australian Science Archives Project launched Bright Sparcs, a web-based biographical database of Australian scientists.⁴ The data in Bright Sparcs was drawn from the Register of the Archives of Science in Australia, and included links to any known archival sources. It was the first online guide to Australian archival holdings.

Bright Sparcs and its successor, the Encyclopaedia of Australian Science, used the technology of the web to emphasise the relationships between people, organisations, archives, and other resources. Researchers could simply follow the links to find new connections and sources.⁵ The relatedness of information was at the core of a range of similar research-focused directories and databases developed over the following decade, including AusStage, AustLit, the Dictionary of Australian Artists Online, the Australian Women's Register, Chinese-Australian Historical Images in Australia, and Reason in Revolt.⁶ These sites captured contextual information around individual resources. Discovery took place within the bounds of an intellectual construct, defined by specific research interests, and enriched by expert knowledge.

Bright Sparcs' data model drew upon the Series System, an innovative approach to archival description developed in the Commonwealth Archives Office (later the National Archives of Australia) in the 1960s.⁷ Rather than thinking in hierarchical terms, the Series System documented each record series within a network of relationships. This approach was implemented as the Commonwealth Record Series System (CRS), and in the mid-1980s it was computerised with the creation of the RINSE database. A separate database, ANGAM II, stored item-level descriptions and allowed keyword searching of file titles. These two systems were combined in 1999 to create RecordSearch, the Archives' online collection access system.⁸

Originally, RecordSearch did not offer a 'simple', one-stop, search box. While item keyword searches were possible, researchers were encouraged to explore the provenance of records – identifying series and agencies related to their topic, rather than going straight to a list of items. One reason for this is that a large proportion of the Archives holdings are not described at item level. I recently harvested a snapshot of series data from RecordSearch. Holdings are recorded for 40,568 series, but almost half of these have no item level descriptions. A further 25 per cent of series are partially described. The 2020 Tune Review noted that only 37 per cent of items are 'described and made discoverable online'.⁹ By my estimate, that means that 23 million items are invisible to keyword searches. But our expectations of search interfaces have been shaped by the power of Google. We expect search to *just work*. Despite its limitations, RecordSearch's default interface was changed to a familiar keyword search box.

Perhaps the demise of the Argus Index should be examined in the context of these sorts of tensions – balancing the role of expert, contextual knowledge against the benefits of simplicity. John Hirst, the leader of the Argus Index Project, had no doubt that the project's carefully curated entry points remained valuable even with the advent of keyword searching. 'You could search "rabbits" by word, certainly, and generate thousands of hits and go down many burrows', he told the Royal Historical Society of Victoria in 2013, 'Or

you could read our entry and see immediately what was going to be useful'.¹⁰ The index offered a 'first sorting' of results to help researchers find what was most relevant to their needs. It also embedded an understanding of changes in language and terminology. 'Searching an old newspaper by word will not work if you do not know what words to use', Hirst argued.

The Tung Wah Newspaper Index tackled an even more complex problem, creating an English-language index to two Chinese-language newspapers, the *Tung Wah News* (1898-1902) and the *Tung Wah Times* (1902-1936). The index was made available online in 2000, opening access to two of 'the most important publications in the history of Australia's Chinese communities'.¹¹ It was not until 2016 that digital copies of these newspapers started to be added to Trove. Like other digitised newspapers, Trove made the text of the *Tung Wah News* and *Times* searchable using Optical Character Recognition (OCR), but for keyword searches to work you need use the original language. There are now at least 51 newspapers in Trove with non-English content, published in a variety of languages and scripts. How discoverable are they? Not only do you need to know what words to use, you need to know what language. Searching only for English keywords will render many of these newspapers invisible. As with RecordSearch you need to start elsewhere, exploring the full range of available titles.

But indexes and finding aids also limit our explorations to topics deemed interesting and sources judged worthy. By doing away with hierarchies of knowledge and value, keyword searching of text offers glimpses into the life of the past that are not mediated by experts. Instead of focusing on 'newspapers of record', Trove opens up more than 1,600 newspapers – small and large, local and national. Trove's newspapers come from across the country, capturing the experiences of many and varied communities, and reflecting different flavours of political and religious thought. Kate Bagnall has described how searching Trove's digitised newspapers has changed her research into Chinese women in Australia – fragments of information from small country newspapers can help build a richer and more inclusive portrait of Australia's past.¹²

Keyword searching of full text resources is a basic research method, but rarely are its limits and biases discussed. Meanwhile, more structured approaches to resource discovery struggle to remain current. All of the specialised databases I mentioned above have had difficulty finding ongoing funding, and some are no longer being updated. It was planned that the Argus Index would be made available as PDFs with links to articles in Trove, but I can find no evidence that this ever happened. The Tung Wah Index was rebuilt in 2014 and page references were later linked to Trove, but you need to know it exists.¹³ This patchwork of half-forgotten links is simply the reality of discovery. While there are ways we might address sustainability and improve integration of resources, dead ends will remain. Navigating these dead ends, working around the limits of search, are now fundamental to historical research online. We need to balance the excitement of full text discovery with an understanding that the systems we use to find, access, organise, and document historical sources are disjointed and fragile.

Digitisation priorities

In 2001, the National Archives of Australia launched a new, highly innovative service. Having investigated the processes, technologies, and costs of collection digitisation, they simply asked users what they wanted. By clicking a button, researchers could submit a particular file to be digitised – at no cost. Despite little publicity, demand was ‘instantaneous’. Between April 2001 and the end of 2003, 77,000 records were digitised, generating 3.5 million images. As Ted Ling and Anne McLean noted, this new ‘digitisation on demand’ service was giving researchers ‘*exactly* what they want’: ‘Our researchers are telling us *precisely* which records are of value to them, and we are able to provide them no matter where the researchers are’.¹⁴ These happy times would not last. The service was originally limited to records held in Canberra. In 2007 it was extended across Australia, but the new ‘national digitisation service’ charged modest fees. Between 2007 and now the cost of digitising a file containing between 100 and 400 pages has increased by more than 500 per cent. A service aimed at delivering access across Australia, became focused on cost recovery.

If access to digitised collections of primary sources has changed the way we do history, then decisions about what gets digitised and when are critical. Who makes these decisions, and how do they shape the types of history we do? In the mid 1990s, academics and librarians worked together on the Australian Cooperative Digitisation Project 1840–45, identifying a set of resources that would both serve as a useful test of digitisation processes, and also open a significant period of Australian history to further examination.¹⁵ The idea that the research community should play an active role in the setting of digitisation priorities was taken up in 2008 by the Australian Academy of the Humanities. Their report on the establishment of an ‘Australian Humanities Digital Archive’ proposed it should be governed by a board comprising representatives of ‘the humanities research community, the e-humanities research community, and the collecting institutions’.¹⁶ That same year, the Commonwealth government’s *Strategic Roadmap for Australian Research Infrastructure* highlighted the role of digitisation in supporting humanities research. A centrally-funded digitisation program would, the Roadmap noted, ‘be characterised by a strategic approach’, with priorities set in consultation with high-level reference groups.¹⁷ Predictably, nothing happened. Without Commonwealth infrastructure funding, GLAM institutions simply got on with digitising what they could, when they could.

Early digitisation efforts often focused on ‘foundational’ collections, reflecting traditional narratives of Australian history that were dominated by explorers and political worthies. These included collections relating to Cook, Banks, Flinders, Barton, Deakin, and Curtin. Notable names also helped attract much-needed digitisation dollars. The Flinders project at the State Library of NSW was supported by James Fairfax, while Dick and Pip Smith kicked in \$50,000 to help get the Barton papers online at the NLA.¹⁸ The works of ‘great men’ remain a marketable commodity as cash-strapped institutions seek support for online projects. The 2019 Federal Budget allocated \$10 million to the NLA to establish a ‘Digitisation Fund’ – its ‘flagship project’ was announced to be the ‘digitisation of the papers of eminent Australians including Sir John Monash and Sir Robert Menzies’.¹⁹

GLAM institutions set digitisation priorities by weighing up factors such as uniqueness, preservation needs, the level of user interest, and the restrictions of copyright. And, of course, they also have to be aware of potential funding sources. But these individual decisions add up to something more – piece by piece they build a representation of the past, only glimpsed in slivers through search results. How do we know what we're looking at? Priorities for the digitisation of Trove's newspapers have been developed through collaboration between the state and national libraries, with user feedback and other community interests factored into calculations.²⁰ In the lead up to the centenary of World War I, for example, it was decided to focus digitisation on the period between 1914 and 1919, particularly in Victoria. Not an unreasonable decision, but one that has significantly skewed the overall distribution of newspaper articles in Trove.²¹ Similarly, the vagaries of copyright law forced a focus on newspapers published before 1955, severely limiting the scope of historical inquiry. But there are exceptions. The *Canberra Times* is available through to 1995 as the result of a special agreement with the publisher. How do these limitations, and agreements affect the way historians use Trove? Will the *Canberra Times* become the 'newspaper of record' for recent history simply because of its online availability? Preliminary analysis of footnotes in articles published by *Australian Historical Studies* indicates that references to post-1954 *Canberra Times* articles have increased significantly since it was digitised.

Collections that show high levels of use, often by family historians, are common targets for digitisation. Genealogy is now big business, and a number of Australian GLAM organisations have entered into agreements with family history sites to digitise relevant records. While these agreements inject much needed dollars into digitisation projects, they can impose new limits on access. The NAA prioritises heavily-used series as part of its 'proactive' digitisation program. The impact of this on overall access to the collection is dramatic. The top five per cent of series, by number of items digitised, contain over 96 per cent of the 2.2 million digitised items in RecordSearch. What are they? The top ten series are all either photographs or military records. Number one on the charts is series B2455, containing 376,000 digitised WWI service records. Before long it is likely to be overtaken by the corresponding WWII series, B883. Digitisation of both these series has been funded by special government allocations, reflecting how the power of the Anzac legend lingers on in official versions of the past. Digitising heavily used collections certainly increases public access, and frees up an institution's resources to respond to other user needs. But there is a danger that such decisions might become self-reinforcing – heavily-used collections are used ever more frequently, simply because they are most easily accessible. Our visions of the past are turned inwards, towards what is familiar and accepted.

The NAA's original digitisation on demand service supported specific research needs alongside bulk digitisation efforts. But the system is now badly out of balance. The latest fee increase in 2016 resulted in a 50 per cent drop in digitisation requests. Instead of helping to build a diverse national collection servicing a wide range of research topics, researchers now assemble their own individual, isolated collections of digital photographs. Their efforts and interests are lost to the commons. Can we find other ways of influencing digitisation priorities? Trove's partner digitisation scheme, while requiring significant investment, does at least allow community organisations to intervene in the setting of priorities by raising

their own funds. For example, the ‘Digitise the Dawn’ campaign was successful in adding Louisa Lawson’s pioneering publication to Trove.²² ARC funding has enabled research projects like ‘Founders and Survivors’ to work with collecting institutions on the digitisation and transcription of selected records, but there is little support for individuals wanting to digitise and share resources related to an original research topic.²³

Platforms for research

What does it *mean* when your search for newspaper articles in Trove returns three million results? The sheer scale of digital resources like Trove challenges our ability to understand, but it also offers possibilities for exploring history in new ways. Ten years ago I created a tool that enabled you to visualise the results of Trove newspaper searches over time.²⁴ Instead of a list of twenty search results, you saw *everything* – the peaks and troughs pointed to changes in language, technology, and the impact of specific events. Even then, Trove was not just a website, or a collection of digitised newspapers, but a source for data for large scale analysis. The original version of this tool, later known as QueryPic, extracted data from the Trove web interface. It was updated a year or so later, when Trove released an API (Application Programming Interface) that made its data directly available for use in new applications. The latest incarnation of QueryPic is available through the GLAM Workbench, along with a suite of other tools and examples that help you work with data from Trove and other digital collections.²⁵

Over the years, QueryPic and related tools have been cited a number of times in the historical literature. Together with a wide range of other tools, packages, directories, and datasets, often created and maintained by volunteers, they make it possible to interrogate digital cultural collections in new ways. However, when it comes to designing national ‘research infrastructure’ to support the humanities in Australia, funding bodies and peak advisory groups remain fixated on big solutions and integrated platforms. The Commonwealth Government’s 2008 *Strategic Roadmap for Australian Research Infrastructure* argued for ‘purpose-built, accessible, interoperable and dedicated HASS eResearch infrastructure’ that would ‘transform current research practices’.²⁶ The 2016 edition hoped ‘to accelerate the impact of HASS research through a single platform that will make dispersed data sets more easily discoverable and accessible’.²⁷

Meanwhile, innovative projects like Australian Museums Online (AMOL) came and went. The National Online Archival Network failed to get beyond the planning stage. The Museum Metadata Exchange stopped being updated. Cultural heritage APIs were created, then quietly decommissioned. Data management tools like OHRM and Heurist struggled to maintain funding.²⁸ Trove certainly succeeded in bringing together a wide range of digital collections and databases including the Australian National Bibliographic Database, Picture Australia, the Register of Archives and Manuscripts, museum collections large and small, and university research repositories. But any user of Trove will know that the aggregated data is often inconsistent, incomplete, or confusing. Nonetheless, with its aggregated search, digitised primary sources, and API, Trove is the closest thing we have to national infrastructure for research in the humanities. So why does it struggle to attract consistent funding from government research coffers?

Integration does not have to be forced from the top. Small improvements can be made to existing resources to better support research practice and open collection data to new forms of analysis. Embedding item metadata within web pages, for example, enables users to assemble their own research data collections using tools like Zotero.²⁹ The 2020 Trove update actually removed access to much item metadata and broke the existing Zotero integration. The importance of persistent links to collection resources cannot be emphasised enough. Yet, RecordSearch still makes it difficult to save and share links to resources, forcing users into complex workarounds. Full text from OCR or transcription should be downloadable. High resolution copies of images should be available for viewing and download. The State Libraries of NSW and Victoria have already started to implement the International Image Interoperability Framework (IIIF) for the online delivery of collection images. Widespread support of IIIF would make it possible for researchers to compare images across collections, and employ a variety of open source tools for analysis and annotation.³⁰ Integrating tools like Hypothes.is into existing interfaces would allow students and researchers to create and share annotated versions of historical resources.³¹ There is nothing risky or radical here – these are well-established tools and technologies.

Alongside these sorts of interface enhancements and integrations, GLAM organisations should be encouraged to make more of their collection data available in machine-readable forms, either through APIs or as downloadable datasets. For example, NSW State Archives & Records holds huge amounts of valuable data in its online indexes. When I last harvested this data in July 2019, there were 1,499,259 rows of data in 64 indexes.³² But most of these indexes are not currently available through the NSW government data portal, and changes to the State Archives site in recent years has made it impossible to access complete versions of the underlying datasets. Neither historians, nor GLAM organisations can avoid the challenges of data for long. Histories of the 1990s or beyond will need to make use of web archives. Increasingly, personal and government archives will be born digital – the comforts and pleasures of paper-based records will give way to new digital interfaces. Greater access to collection data will power new perspectives on Australian history, allowing researchers to work across multiple scales, bridge institutional silos, find new patterns and connections, and ask new questions of the past.

All of this takes money. Current infrastructure funding schemes are focused on research institutions themselves, and favour large unwieldy projects with many institutional partners. Where is the funding for GLAM organisations wanting to open their datasets, or build new access points for researchers? Where is the funding for maintaining existing platforms and databases? Where is the funding for innovative, lightweight, digital tools that meet specific needs of researchers? By focusing on targeted, small-scale interventions, rather than woolly visions of national integration, we can build an infrastructure framework that both responds to researchers and rewards experimentation.

But there is no need to wait. We have the tools and data to make a start. We need to stop being distracted by government priorities or the latest tech coolness and start using what we have. Use drives innovation. Use sets priorities. Use justifies investment. The question then becomes, how do we support historians in exploring the broader possibilities of digital collections? Resources like The Programming Historian and the GLAM Workbench

provide useful starting points, however, the development of digital skills takes time.³³ From undergraduate methods courses, through to postgraduate research training, and professional development workshops, the ability to use and critique digital collections needs to be recognised simply as historical practice. In the end it is not a matter of technology, but of valuing the work involved in writing tutorials, documenting datasets, sharing code, and building tools. This too is history.

Gains and losses

Just twenty years ago, historical research often entailed long hours spent at a microfilm reader, browsing newspaper after newspaper in the hope of finding something relevant. The changes wrought by Trove, and other digital collections, seem revolutionary, but as with all revolutions there have been gains and losses. Alongside the wonders wrought by digitisation, this chapter has tried to highlight some of the paths not taken. The online resources we now use daily are not simply the products of technology – priorities have been set, funding has been distributed, decisions made about what to include and what to leave out. Cultural heritage collections are not just put online, they are placed within specific contexts of discovery and use. Each object, each version, each interface comes with a set of limitations and affordances that together determine what is possible. We do not know yet how these decisions will shape the sorts of histories that we write.

In 2013, Tim Hitchcock noted that while keyword searching of historical texts had been ‘fantastically liberating’, it had also uprooted ‘the components of what was once a coherent collection of beliefs and systems for discovering and performing taxonomies on information’.³⁴ Historians, he argued, needed to engage more fully with new technologies of discovery: ‘At the moment we are using them to make our lives easier, while pretending that they do not exist’.

Similarly, the construction of ‘access’ has been rendered invisible by the power and novelty of technology – it just works. The critical analysis that historians bring to the context and meaning of any primary source now need to be applied to the digital tools that deliver those sources to our desktops. The revolutions will continue, and we need always to ask, who wins and who loses.

¹ John Hirst, ‘On indexing the Argus’, in *Victorian Historical Journal*, vol 84, no 1, June 2013, pp 6–15.

² The Argus Index Project (Online), 2009. Available: <https://web.archive.org/web/20091113123141/http://www.nla.gov.au/apps/argus?acton=Menu&type=about> (Accessed 28 May 2021)

³ ‘Argus On-Line: A Nineteenth Century Australian Newspaper Digital Index’, LE0453714, ARC Grants Database (Online). Available: <https://dataportal.arc.gov.au/NCGP/Web/Grant/Grant/LE0453714> (Accessed 28 May 2021).

⁴ Ailie Smith and Gavan McCarthy, 'The Encyclopedia of Australian Science: a virtual meeting of archives and libraries', *The Australian Library Journal*, vol 65, no 3, 2 July 2016, pp 191–202. Available: <https://doi.org/10.1080/00049670.2016.1212318> (Accessed 28 May 2021).

⁵ Encyclopaedia of Australian Science (Online). Available: <https://www.eoas.info/> (Accessed 7 June 2021).

⁶ AusStage: The Australian Live Performance Database (Online). Available: <https://www.ausstage.edu.au/> (Accessed 7 June 2021); AustLit (Online). Available: <https://www.austlit.edu.au/> (Accessed 7 June 2021); Design & Art Australia Online (formerly Dictionary of Australian Artists Online). Available: <https://www.daaao.org.au/> (Accessed 7 June 2021); Australian Women's Register (Online). Available: <https://www.womenaustralia.info/> (Accessed 7 June 2021); Chinese-Australian Historical Images in Australia (CHIA) (Online). Available: <https://www.chia.chinesemuseum.com.au/> (Accessed 7 June 2021); Reason in Revolt (Online). Available: <https://www.reasoninrevolt.net.au/> (Accessed 7 June 2021).

⁷ Mark Wagland and Russell Kelly, 'The Series System – A Revolution in Archival Control', in Sue McKemmish and Michael Piggott (eds), *The Records Continuum: Ian Maclean and Australian Archives first fifty years*, Ancora Press, Monash University, 1994, pp 131–149.

⁸ Paul Macpherson, 'Improving Access for the Public to the Collection of the National Archives of Australia', *Australian Academic & Research Libraries*, vol 31, no 2, January 2000, pp 79–90. Available: <https://doi.org/10.1080/00048623.2000.10755118> (Accessed 8 May 2021).

⁹ 'Functional and Efficiency Review of the National Archives of Australia (Tune Review)', Attorney-General's Department, 2021. Available: <https://www.ag.gov.au/rights-and-protections/publications/tune-review> (Accessed 8 May 2021).

¹⁰ Hirst, 'On indexing the Argus', pp 8–9.

¹¹ The Tung Wah Times: A window into Chinese community history (Online). Available: <https://arrow.latrobe.edu.au/store/3/4/5/5/1/public/tungwah.htm> (Accessed 7 June 2021).

¹² Kate Bagnall, 'Chinese women in colonial New South Wales: From absence to presence', in *Australian Journal of Biography and History*, no 3, 2020. Available: <http://doi.org/10.22459/AJBH.2020.01> (Accessed 29 May 2021).

¹³ Tung Wah Newspaper Index (Online). Available: <https://resources.chineseaustralia.org/tungwah/> (Accessed 7 June 2021).

¹⁴ Ted Ling and Anne Mclean, 'Taking it to the People: Why the National Archives of Australia Embraced Digitisation on Demand', in *Australian Academic & Research Libraries*, vol 35, no 1, January 2004, pp 2–15. Available: <https://doi.org/10.1080/00048623.2004.10755253> (Accessed 8 May 2021).

¹⁵ Ross Coleman and Alan Ventress, 'The Australian Cooperative Digitisation Project 1840-45 - A Progress Report', in *LASIE*, vol 28, no 1, March 1997, pp 46–50. Available:

<http://hdl.handle.net/2123/1425> (Accessed 30 April 2021); Ross Coleman and Colin Webb, 'Digital conversion of Nineteenth century publications - production management in the Australian Cooperative Digitisation Project 1840-45', in *LASIE*, vol 31, no 2, June 2000, pp 5–20. Available: <http://hdl.handle.net/2123/1424> (Accessed 30 May 2021).

¹⁶ Graeme Turner, 'Towards an Australian Humanities Digital Archive', Australian Academy of the Humanities, 2008. Available: <http://www.humanities.org.au/wp-content/uploads/2017/04/AAH-Digital-Humanities-Archive-2008.pdf> (Accessed 15 May 2021).

¹⁷ *Strategic Roadmap for Australian Research Infrastructure*, Department of Innovation, Industry, Science and Research, Canberra, 2008. Available: <https://www.dese.gov.au/national-research-infrastructure/resources/national-collaborative-research-infrastructure-strategy-strategic-roadmap-2008> (Accessed 12 May 2021).

¹⁸ Warwick Hirst and Paul Scifleet, 'Voyaging with Matthew Flinders: The James Fairfax Matthew Flinders Electronic Archive', in *LASIE: Library Automated Systems Information Exchange*, vol 31, no 1, March 2000, pp 24–27; Andrew Stawowczyk Long, 2001. Now they know my name: creating a digital presence for Barton (Online). Available: <https://ses.library.usyd.edu.au/handle/2123/6224> (Accessed 7 June 2021).

¹⁹ ALIA, 2020. Federal Government Budget 2019-2020 (Online). Available: <https://www.alia.org.au/news/18411/federal-government-budget-2019-2020> (Accessed 7 June 2021).

²⁰ For more on the construction of Trove's newspaper corpus, see: Katherine Bode, *A World of Fiction: Digital Collections and the Future of Literary History*, University of Michigan Press, 2018, pp 67-73. Available: <https://hdl.handle.net/2027/fulcrum.5q47rp73f> (Accessed 13 May 2021); M H Beals and Emily Bell 2020, *The Atlas of Digitised Newspapers and Metadata: Reports from Oceanic Exchanges* (Online). Available: <http://www.doi.org/10.6084/m9.figshare.11560059> (Accessed 7 June 2021).

²¹ Tim Sherratt, 'Hacking heritage: understanding the limits of online access', in Hannah Lewi et al (eds), *The Routledge International Handbook of New Digital Practices in Galleries, Libraries, Archives, Museums and Heritage Sites*, Routledge, 2020, pp 116–130. Available: <https://timsherratt.org/blog/hacking-heritage/> (Accessed 7 June 2021); GLAM Workbench: Trove newspapers (Online). Available: <https://glam-workbench.net/trove-newspapers/> (Accessed 7 June 2021).

²² Digitise the Dawn (Online). Available: <https://digitisethedawn.org/> (Accessed 7 June 2021).

²³ Founders & Survivors (Online). Available: <https://foundersandsurvivors.com/> (Accessed 7 June 2021).

²⁴ Tim Sherratt 2011, *Mining the Treasures of Trove (Part 2)* (Online). Available: <http://discontents.com.au/mining-the-treasures-of-trove-part-2/> (Accessed 1 June 2021)

²⁵ Tim Sherratt, GLAM Workbench (Online). Available: <https://glam-workbench.net/> (Accessed 7 June 2021).

²⁶ *Strategic Roadmap for Australian Research Infrastructure*, 2008, p 39.

²⁷ *2016 National Research Infrastructure Roadmap*, Department of Education and Training, 2017. Available: <https://docs.education.gov.au/node/43736> (Accessed 12 May 2021).

²⁸ Heurist (Online). Available: <https://heuristnetwork.org/> (Accessed 7 June 2021).

²⁹ Zotero (Online). Available: <https://zotero.org/> (Accessed 7 June 2021).

³⁰ IIF: International Image Interoperability Framework (Online). Available: <https://iif.io/> (Accessed 7 June 2021).

³¹ Hypothes.is (Online). Available: <https://web.hypothes.is/> (Accessed 7 June 2021); for example, Tim Sherratt 2018, An experiment in two-way direct linking using Hypothes.is (Online). Available: <https://timsherratt.org/research-notebook/historic-hansard/notes/two-way-direct-linking/> (Accessed 7 June 2021).

³² GLAM Workbench: NSW State Archives (Online). Available: <https://glam-workbench.net/nsw-state-archives/> (Accessed 7 June 2021)

³³ The Programming Historian (Online). Available: <https://programminghistorian.org/> (Accessed 7 June 2021); Tim Sherratt, GLAM Workbench (Online). Available: <https://glam-workbench.net/> (Accessed 7 June 2021).

³⁴ Tim Hitchcock, 'Confronting the Digital', in *Cultural and Social History*, vol 10, no 1, 1 March 2013, pp 9–23. Available: <https://doi.org/10.2752/147800413X13515292098070> (Accessed 29 May 2021).