



# DISCONTENTS



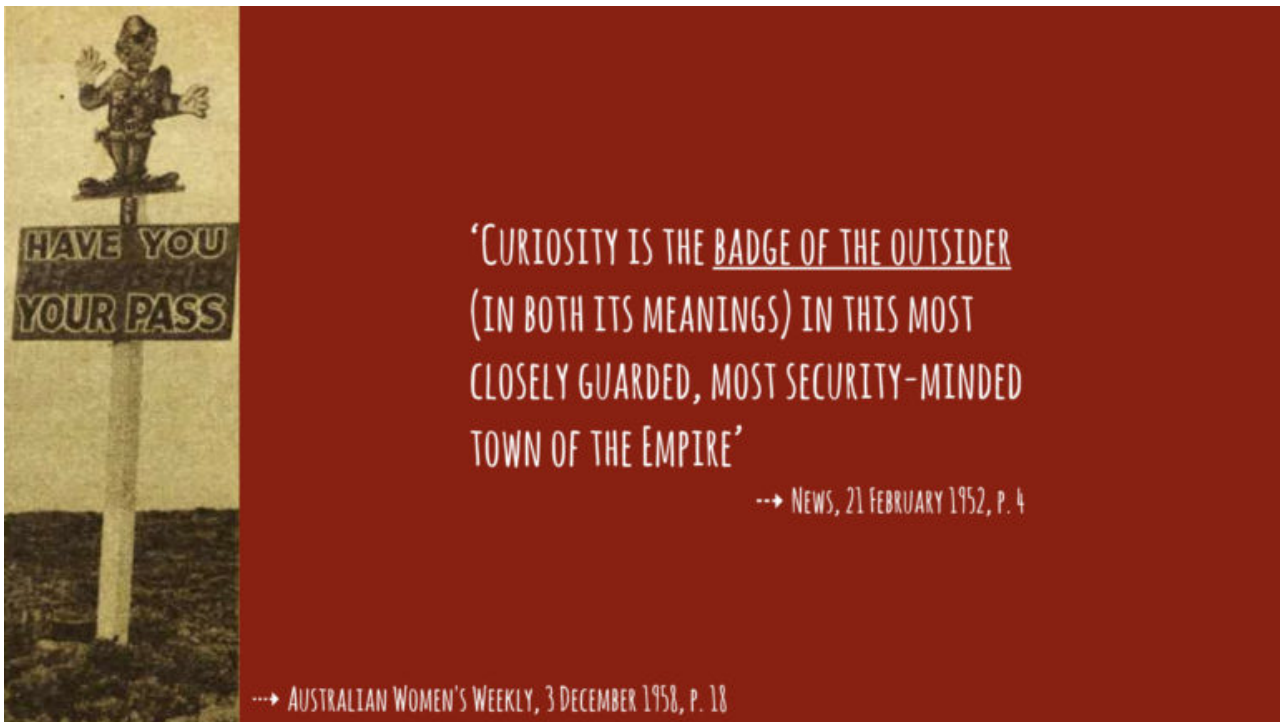
NOVEMBER 24, 2017 / TIM SHERRATT

## 'THE BADGE OF THE OUTSIDER': OPEN ACCESS AND CLOSED BOUNDARIES

Presented at [\*Sharing is Caring 2017\*](#), 20 November 2017, in Aarhus, Denmark.  
You can also [watch the video](#).

---

In 1946, Britain decided that Australia would be the perfect place to test missiles. The Australian government, keen to play its part in the defence of the Empire, readily agreed. Ignoring, yet again, the presence of Australia's Indigenous peoples, defence planners thought Australia was attractive to because it was 'empty', flat, and far from 'prying eyes'.



The town of Woomera was built in the South Australian desert to house scientists, workers, and military personnel. It was a town where no housewife could go to the shop without her security pass; where curiosity was 'the badge of the outsider'.

But while Australia's land seemed ideal for secret military operations, its people remained suspect. Britain's plans were threatened by concerns about Communist infiltration of the Australian government. Under pressure from the UK and USA, Australia sought to lift its spy game through the establishment of a new agency to monitor such threats — the Australian Security Intelligence Organisation, or ASIO.

Legislation defined ASIO's functions in very broad terms, 'to obtain, correlate and evaluate intelligence relevant to security'. From the 1950s to the 1970s, this was used to justify surveillance of a wide range of potential 'subversives' — not just known Communists, but writers, artists, academics, scientists, Indigenous activists, and more. Many thousands of files were created to document their beliefs, activities, connections, and personal lives. Recordkeeping was critical to the practice of state surveillance.

We don't know how many files were kept on ordinary Australians because ASIO is exempt from many of the key provisions of the government's archives legislation. Unlike other agencies, ASIO does not routinely transfer records or indexes to the National Archives of Australia. Researchers have to go on a fishing expedition, asking the National Archives to ask ASIO whether they might have a file relating to a particular person or organisation. If ASIO admits it has a relevant file in the open period (more than twenty years old), the file goes through an 'access examination' process to determine whether it contains information that should be withheld for reasons of

national security, or individual privacy. If anything is left, it is finally opened to public access.

Despite these hurdles, more than 12,000 ASIO surveillance files have been made public, though most include redactions — black boxes obscure words too sensitive to be read.

The files have been used in biographies, family histories, and studies of Australia's literary community. One recent book invited the subjects of ASIO surveillance to reflect on the contents of their own files — to see their lives through a different set of eyes; to explore the intrusions and innuendo that passed for 'intelligence'.

I'm currently working with a set of 60,000 photographs held by the State Library of New South Wales. These photos were taken for *The Tribune*, a Communist Party newspaper published in Sydney, and document protest and political activity in Australia from the 1960s to the 1990s. One of the things I'm interested in is finding overlaps between the Tribune photographs and ASIO surveillance files. For example, in February 1972 there was a demonstration on Indigenous rights held outside Parliament House. Because both sets of records have been digitised and made public, we can compare perspectives — spies versus 'subversives'.



Tribune photograph of speakers at the Aborigines' (sic) surveillance photograph of the same event



**Photo Credits:** **Before** State Library of NSW: Tribune negatives, Series 1, Item 680 **After** National Archives of Australia A9626, 315, p. 8

This is a reminder that the impact of digitisation is not simply easier, more immediate, access. We can also see the same things differently. We can interrogate the meaning of access itself.

---

RecordSearch, the National Archives of Australia's online database, provides access to about 64,000 series descriptions, 11 million item descriptions, and 1.8 million digitised pages. There's currently no API, or downloadable datasets at item level, so I make my own.

For six or seven years now I've relied on my own little library of screen scrapers to get data out of RecordSearch. They're slow and they break easily, but they do the job.

Late last year I embarked upon what was probably my most ambitious data harvest. I gathered information about every series listed on RecordSearch and calculated, for each, the quantity of records (in linear metres), the number of individual items described, and the number of items digitised. I then aggregated the series by the top-level functions of agencies associated with them. Basically I grouped them by subject — defence, security, education etc.

Why?

Because digitisation shapes our perceptions of reality. The more we have in digital form, the easier cultural heritage collections are to find and use, the more likely we are to assume that everything (or at least everything important) is online. Ease of access bears an ontological weight — if we can't find it online, does it exist?

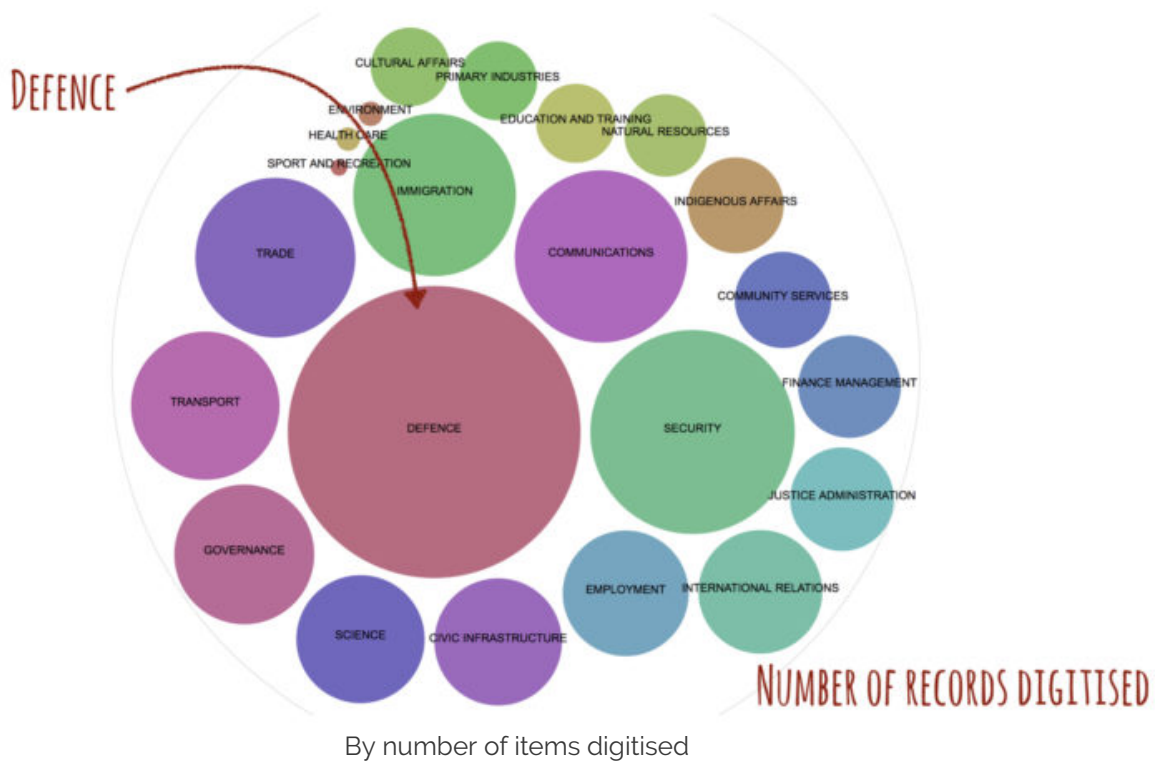
Now that might not be a problem if what was digitised somehow provided a representative sample of the whole. But we all know how such decisions are shaped by political priorities, funding opportunities, user demand, public events, and happy accidents. There's nothing necessarily wrong with that, it's just the environment within which we work. There are never enough resources. We have to do what we can, when we can.

The problem is, we rarely expose the impact of these decisions to the users of our digital collections. We rarely give them the chance to reflect on how our decisions shape their assumptions.

The National Archives of Australia documents the workings of our democracy. It offers one important perspective on who we are as a nation. If we look at the quantity of records associated with each top-level function, we see a fairly even distribution. Nothing stands out.



But what happens when we view the activities of government through the number of files digitised in each subject area?



The prominence of defence is really no surprise. Service records are heavily used by family historians, and in 2007 the Australian government funded the digitisation of all 375,000 World War I service records in what was branded as 'A Gift to the Nation'.

The National Archives is not alone. I often show people this [graph of the number of digitised newspaper articles](#) in Trove, pointing out the fairly dramatic peak around 1914.

Did something happen in 1914? Were there more articles published, more newspapers? No, there's just more money. In the lead up to the centenary of WWI, funding was directed towards the digitisation of newspapers from the wartime period.

Again, there's nothing wrong with this. It's just that these biases are not obvious to someone typing queries into a search box. In the context of Australian history, these decisions around digitisation help to reinforce the long-held belief that Australian national identity was somehow forged on the battlefields of WWI. It helps to put war at the centre of our history, at the centre of who we are.

But of course while digitisation can shape our assumptions, it also gives us new opportunities to critique them. I could only analyse the holdings of the National Archives because their collection data is online. We don't have to take just what the search box delivers — we can ask our own questions. But this is only possible if people have the skills, the tools, and the confidence to poke around in the data. This too is access. Institutions should invite the public not to swoon at their digital delights, but to hack away at difficult questions — not to see collections, but to see them in unexpected and challenging ways.

---

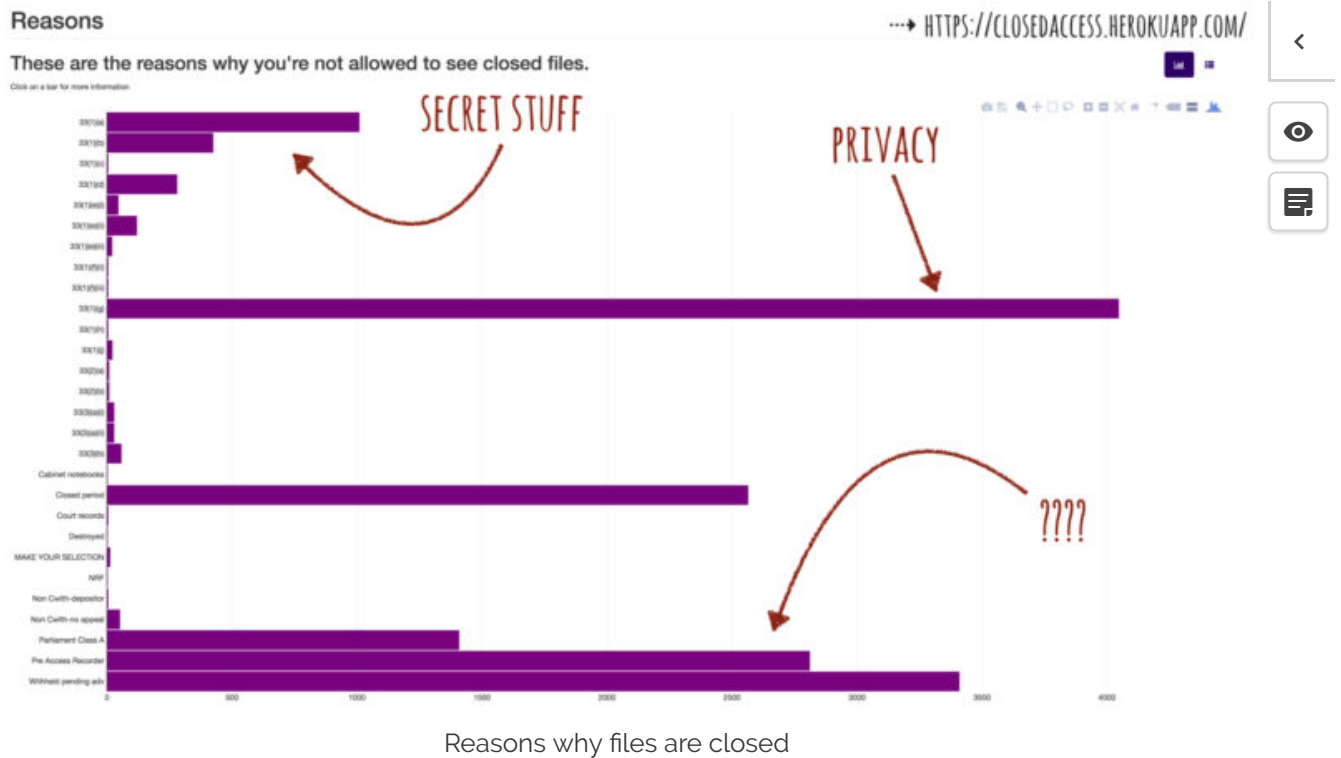
I mentioned that ASIO files go through an process known as 'access examination' before they're released to the public. This is the case for all records more than twenty years old, not just the super secret ones. The vast majority of files are simply opened without restriction. Some, including most of the ASIO files, are opened 'with exceptions' — pages can be withheld, and text redacted. A few are withheld from the public completely. They have entries in RecordSearch, but you can't see them — their access status is officially 'closed'.

But because the metadata about access decisions is available online, we can start to build a picture of what we're not allowed to see.

At the start of 2016, I harvested the details of all files in the National Archives of Australia with the access status of 'closed'. I've aggregated and sliced the data in a number of different ways, so you can explore the age of the files, what series they came from, and when decisions were made about their access status. At any point you can drill down to a list of the files you cannot see — making it perhaps the most frustrating search interface ever devised.







You can also examine the reasons why files have been withheld. Many of these exceptions are defined by the legislation that established the National Archives. Clause 33(1)(a), for example, relates to national security, 33(1)(g) is concerned with individual privacy. But the metadata reveals that files are withheld for a number of other reasons, such as 'Pre Access Recorder' and 'Withheld Pending Advice'. There's also, you might note, a category entitled 'MAKE YOUR SELECTION' — which reveals something about the limits of the data entry interface.

By poking around you in the data you can make some guesses as to how these additional categories are used. 'Pre Access Recorder' is used as a catch-all for records that were withheld from public access before the archives legislation was passed. 'Withheld Pending Advice' is used to label files that have been sent off to other government agencies for their assessment — they're not yet finally closed, but as this process can take years, they're sort of closed. Indeed, my interface shows that 1,467 files have been waiting more than three years for advice.

The point of this is not to embarrass the National Archives, nor the Department of Foreign Affairs and Trade which holds the most files in limbo. The point is to examine the ways in which access itself is constructed. Legislation defines an ideal, but the reality is more messy and human. By tracking patterns in the way access decisions are made we can explore the historical processes at work. Access is not allowed, it is made.

Remember those 12,000 ASIO files publicly available through the National Archives? You might not be surprised to know that I've harvested them all — both the metadata



and the 300,000 digitised pages. There's about 70gb of images.

Using these files we can dig a little deeper into the nature of access. I wrote a [computer vision script](#) to find redactions. It took a lot of trial and error, and I'm about to start work on a smarter version that incorporates machine learning, but it did the job. From one series of ASIO files, about 230,000 pages, I extracted 239,000 redactions — lots and lots of little black boxes. You'll be pleased to know that not only can you [download the complete set of redactions](#) from the research repository Figshare, you can [browse them](#). All of them! Hours of fun for all the family!



redacted

The interesting thing about this interface is that if you click on a redaction you can view the page that it was extracted from. So it's sort of an inside out discovery interface. Instead of the redactions being a brick wall or a dead end, they're a starting point. A practice intended to remove information, to limit access, becomes a gateway for exploration. Indeed, the redactions themselves provide an identifiable data point — something that can be analysed to turn the gaze of government surveillance upon itself.

But something else was hiding in those ASIO files. As I was reviewing the collection of redactions for false positives I discovered that someone tasked with the removal of information, decided to add a little creative flair.

I discovered [#redactionart](#).



## The Redaction Zoo

from [Tim Sherratt](#)

03:06

I assure you that these creations really are sitting inside ASIO files held by the National Archives. But since I've discovered them, they've developed a life of their own. Not only can you browse through them online, you can wear them.



I gave away about 80 of these badges at an exhibition earlier this year. To create the badges, I simply traced around the original images and saved the results as SVG files. These files themselves are [shared through GitHub](#) for anyone wanting to create their own #redactionart.



This [amazing #redactionart dress](#) was made by Bonnie Wildie, a librarian in NSW. My SVG files have also been turned into a set of [3D printable cookie cutters](#), as well as a [range of t-shirts and stickers](#) on RedBubble.

This escape from the archives is not only creative and fun, it's important.

It's important because it emphasises that the practices through which government information is controlled and withheld are profoundly human. People make decisions and they leave their marks. There is nothing mysterious or otherworldly in the secret — it is an exercise of power.

---

Archives are not just made of documents — there are people inside.

'Surveillance' is not included in the National Archives' official thesaurus of government functions, yet the movements and activities of individuals are recorded in many thousands of files across an assortment of agencies.

A simple query of my harvested data reveals that the phrase 'alien registration' appears in the titles of only 29 series. But these series contain more than half a million files. 4.7% of digitised files in the National Archives document the movements of so-called 'aliens'. While these registration systems were created during wartime, they lingered beyond. And they were not the only means of keeping track of potential threats. Just as at Woomera, boundaries were drawn, and outsiders marked for attention.



When I was last in this part of the world, I talked a bit about some work that Kate Bagnall and I had done with records of the White Australia Policy held by the National Archives.

A quick recap — when the Australian colonies federated in 1901, it was generally assumed that the new nation's future could only be assured through strict racial homogeneity. A 'white' Australia was a strong Australia. Legislation was quickly passed to restrict immigration and set the foundations for what became known as the White Australia Policy.

However, in 1901 there were around 40,000 people living in Australia whose background was neither European nor Indigenous — they were Chinese, Japanese, Syrian, Indian, and Malay. Some had been born in Australia, or had lived there for many years — raising families, building businesses; just living their lives.

If any of these people wanted to travel overseas they had to carry special documents, or they might not be allowed to return home. Customs officials at Australian ports would ask anyone who seemed not to be 'white' for identification. The badge of the outsider was the colour of their skin.



An example of a Certificate Exempting from Dictation Test, NAA: ST84/1, 1909/21/91-100, p. 35-6

Many thousands of these documents, the remnants of a racist bureaucratic system, are preserved in the National Archives.

Back in 2011, I downloaded about 12,000 of these documents from RecordSearch and ran them through a facial detection script to create a seemingly endless scrolling wall

of faces. We called it '[The Real Face of White Australia](#)'. It's another inside-out interface — instead of showing the files, you see the people inside.



That was then, this is now! In the last few months, I've been working with a group of my [digital heritage students](#) to develop [a website for the collaborative transcription](#) of these same records. We want to put names to the faces. We want to chart their journeys. We want to document their lives.



Our project has no funding, and was only possible because Zooniverse and the New York Public Library created and shared [Scribe](#), a framework for the transcription of structured documents — an easy way to get usable data out of forms, ledgers, and certificates.

The site was launched at a 'transcribe-a-thon' held at the Museum of Australian Democracy in Canberra, which just happens to be located in Australia's first parliament house. The building didn't exist when the Immigration Restriction Act was passed in 1901, but it was where the White Australia Policy was elaborated and maintained.



Busy transcribers at the Museum of Australian Democracy

Transcription continues. There's still much work to do on the documents, but data is already flowing. I'm making regular dumps available for download through a GitHub repository.

But it was never just about the data. Many more people now know that these records, this history, exists. Through the process of transcription you are confronted by the disturbing reality of the records — you're surprised, puzzled, shocked, and often moved. Creating a space for these sorts of experiences is important in itself.

The Museum of Australian Democracy not only gave us their building for a weekend, they let us play with their data projectors. In some ways, I would have been happy if all we had achieved was this — to put these faces in this space.



## The Real Face of White Australia Transcribe-a-thon

from [Tim Sherratt](#)

01:18 |



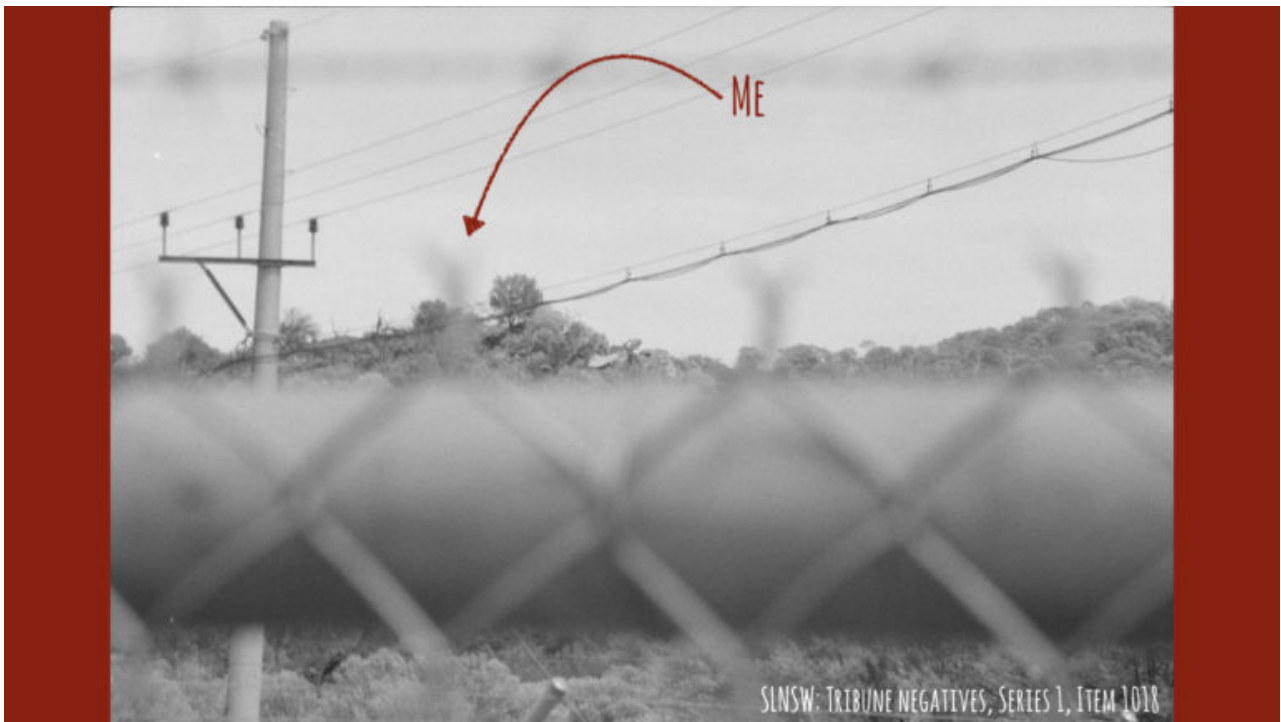
Once again the gaze of surveillance is reversed. In the home of Australian democracy, people who lives were monitored under a racist system of exclusion and control were looking at us, asking questions of us.

Amongst those Tribune photos at the State Library of NSW, I recently found this. Believe it or not, I'm the spy on the right. This compelling piece of street theatre was performed at the gates of Pine Gap, a US electronic surveillance facility right in the centre of Australia. Pine Gap's lease was due for renewal in 1987, so hundreds of protestors converged on the site, hoping that the Australian government might withdraw it's support. Needless to say, it didn't. Pine Gap remains, and in recent times has been implicated in US drone strikes



I found another photo of myself amongst the Tribune archives. A group of us climbed over the outer perimeter fence in the middle of the night and took up positions on a rocky outcrop that overlooked the main gate. At a predetermined time, we leapt out of our hiding places and lit smoke flares. I was arrested soon after, charged with trespass, and fined \$100.



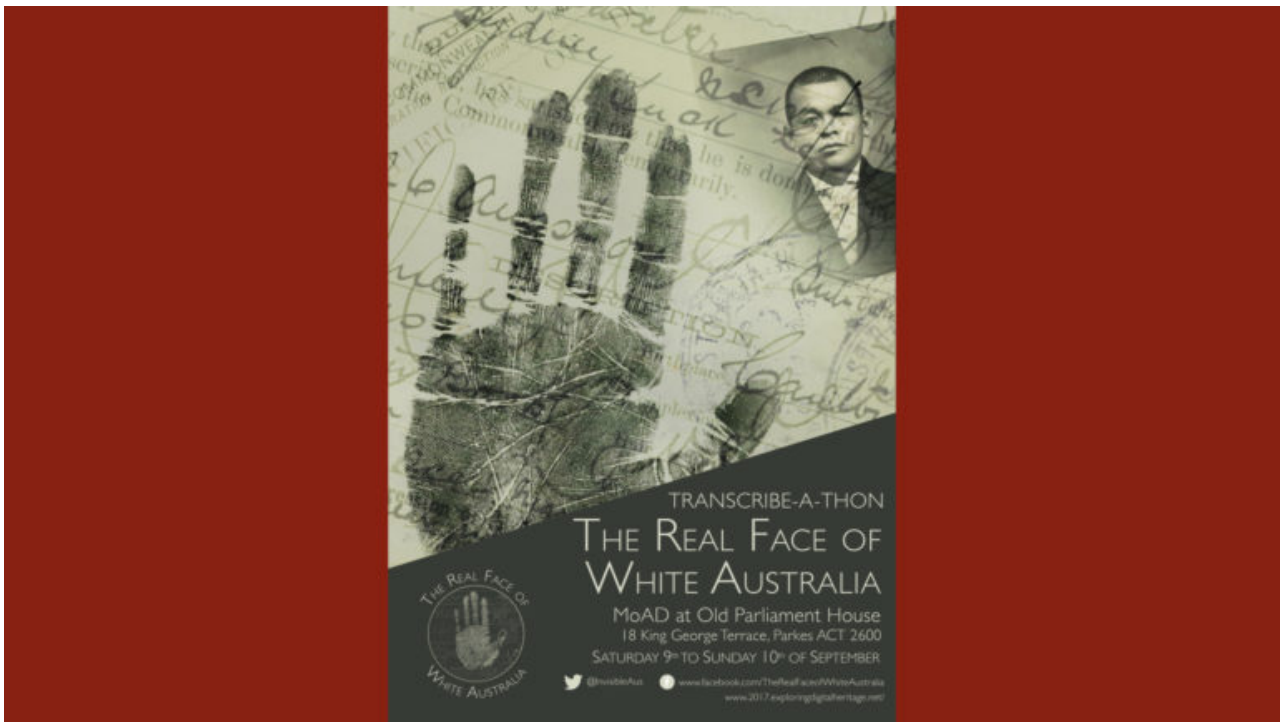


Another group of Pine Gap protestors are currently on trial in Australia. They made it through the protective fences and dared to play music and pray. For this they have been charged under the Defence (Special Undertakings) Act which carries a maximum sentence of seven years in prison. This Act was passed in 1952 when Britain decided to expand its weapons testing program in Australia to include atomic bombs. It expanded upon earlier legislation that had been intended to protect Woomera from Communist interference. This is one of the very few times anyone has been charged under the Act, despite there being hundreds of arrests like mine in the past.

As security services gain new powers, and electronic surveillance expands, it's hard not to see the Pine Gap proceedings as an attempt to discourage criticism of the government's tough on terrorism stance.

At a recent symposium on collaboration between researchers and collecting institutions, Seb Chan described some of the advances that had taken place in opening up collections, but then asked 'So what?'

I suppose that's the question we're hear to discuss. Why do we put all this effort into digitising collections, building interfaces, and sharing data? Easier access is great, beautiful interfaces are cool, but... so what? For me, as a historian, hacker, and sometime heritage professional, the answer is straightforward — it's all about bringing the past into conversation with the present. It's about mobilising our collections as critical resources in debates about who we are, what matters, and why we should care.



Transcribe-a-thon poster designed by Emily Fry


Those inky, black handprints on the White Australia records moved one of my students to reflect on her experience as a recent immigrant from Canada, required by the Australian government to supply a set of her fingerprints. She wrote a beautiful talk and presented it during the transcribe-a-thon in the original House of Representatives chamber at Old Parliament House. Another student noted in her final essay that the documents made non-white residents seem like criminals, pointing to parallels with the current treatment of refugees. On the flip-side, our efforts attracted the attention of a few racist trolls, one of whom referred to the White Australia Policy as 'the good old days'.


Once again 'outsiders' are being targeted as threats to our security. Boundaries are being reinforced, and efforts being made to define who belongs. We know this. We've seen this before. Europeana's new project on the history of migration is an important initiative — we need to tell our stories, share our resources, grapple with our difficult and painful pasts. I don't think this is a time to reassert the authority of our cultural institutions as reservoirs of truth. We are implicated in all of this. Our collections are built upon systems of surveillance, on attempts to put humans into categories. They are products of power and privilege. We are not the guardians of enlightenment, we are the keepers of horrors.

Just like #redactionart, the value of our collections lies in their complexity and contradictions — in their very humanity and all the confusion that entails. Digital collections lend themselves to an exploration of complexity. We can shift scales and perspectives, we can manipulate contexts, we can set collections loose in public spaces, we can turn them inside out. We can see differently, but perhaps more importantly, we can feel differently.

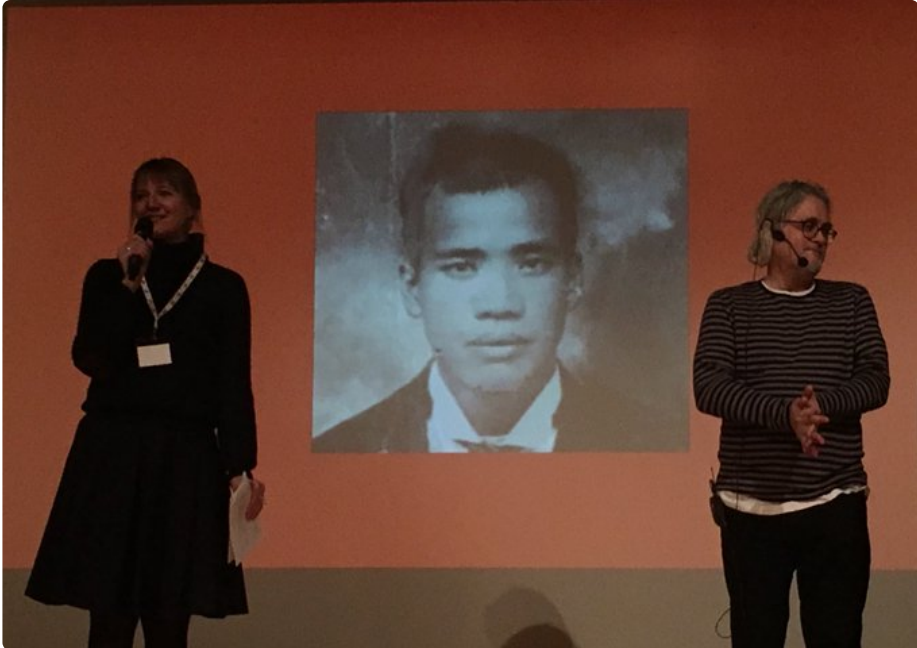
When you think about it, 'impact' is a pretty violent sort of word. There are perhaps a few people around the world we'd like to 'wallop' with our digital collections. But I suspect most of the time we're after something more subtle — to expand possibilities, to undermine assumed certainties, maybe even to expose a glitch in the Matrix.

Perhaps we can offer a glimpse of an alternative reality, where we recognise the outsider as us.



 **Johanna Berg**  
@johannaberg

Maybe in the end we will be able to see the outsider as us. Hear hear! [#ShareCare17](#)



10 10:14 PM - Nov 20, 2017

[See Johanna Berg's other Tweets](#)

[Download PDF](#)

Like this? Take a second to support Tim Sherratt on  
Patreon!

 **BECOME A PATRON**

Share this:



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

---

**PREVIOUS POST**

FINDING A BALANCE

**NEXT POST**

2017 — THE MAKING AND THE TALKING

**CATEGORIES**

ARCHIVES CONFERENCE PRESENTATIONS

**TAGS**

ASIO DIGITISATION RECORDSEARCH REDACTIONS WHITE AUSTRALIA POLICY



**WRITTEN BY:**  
TIM SHERRATT

I'm a historian and hacker who researches the possibilities and politics of digital cultural collections.



**BE FIRST TO COMMENT**

Leave a Reply

Your email address will not be published. Required fields are marked \*



Enter Your Comment...

Your Name\*

Your Email\*

Your URL (optional)

Post Comment

Notify me of new posts by email.

# DISCONTENTS



## ELSEWHERE

TIMSHERRATT.ORG  
@WRAGGE ON TWITTER  
IDENTIFIED BY TROVE  
IDENTIFIED BY ORCID  
RESEARCH LIBRARY AT ZOTERO

CODE AT GITHUB  
PRESENTATIONS AT SLIDESHARE



## SUBSCRIBE

Enter your email address to subscribe to this blog and receive notifications of new posts by email.

Subscribe

## RSS

 [RSS - Posts](#)

TRACKS WORDPRESS THEME BY COMPETE THEMES.