

mostly  
Australian

# EXPLORING GLAM DATA

Tim Sherratt · @wragge · [timsherratt.org](http://timsherratt.org)



THESE SLIDES



<https://slides.com/wragge/resbaz-2021>



## Physics in Australia to 1945



SEARCH

Created by: [Tim Sherratt](#)  
Updated by: [Denise Sutherland](#) and [Elissa Tenkate](#)  
Date Modified: 18 August 1997

AUSTRALIAN SCIENCE  ARCHIVES PROJECT



Published by: Australian Science Archives Project on ASAPWeb  
Comments or questions to: [ASAPWeb \(asapweb@asap.unimelb.edu.au\)](mailto:asapweb@asap.unimelb.edu.au)  
Prepared by: Tim Sheratt  
Updated by: Elissa Tenkate  
Date modified: 30 October 1997

1994-1997



The World-Wide Web Virtual Library - [Alphabetical](#) & [Library of Congress](#).



## History of Science, Technology and Medicine

## Overview

The WWWVL for the history of science, technology and medicine was established on 6 September 1994 to keep track of information facilities in the field of the history of science, technology and medicine. Please mail [Tim.Sherratt@asap.unimelb.edu.au](mailto:Tim.Sherratt@asap.unimelb.edu.au) if you know of any relevant resources not in this list.

[\[ Overview \]](#) [\[ New \]](#) [\[ Alphabetical \]](#) [\[ Lists \]](#) [\[ Organisations \]](#) [\[ Specialised \]](#) [\[ Biographies \]](#) [\[ Institutions \]](#) [\[ Museums \]](#) [\[ E-journals \]](#) [\[ Email \]](#)  
[\[ Search \]](#) [\[ Suggest URL \]](#) [\[ Sign our Visitors' Book! \]](#)

### What's new in HSTM?

[Complete alphabetical listing of resources](#)

[Other lists, collections and bibliographies](#)


## Organisations and Conferences

### Specialised collections and documents

- [Archives and manuscripts](#)
- [Astronomy and Space](#)
- [Australia](#)
- [Biology](#)
- [Chemistry](#)
- [China](#)
- [Communications](#)
- [Computers and Information Processing](#)
- [Early Science and Alchemy](#)

SEEING DIFFERENTLY

# WWI SERVICE RECORDS



RecordSearch

Login

HelpRecordSearch ForumView previous searches

Basic searchAdvanced searchNameSearchPhotoSearchPassenger arrivalsNewly scanned records

Item listWelcome guest

New searchRefine search

Display  search results per page

⏮

⏪






⏩

⏭

Displaying 1 to 20 of 20000

☒ List report

☐ Details report

Select	Series no.	Control symbol	Item title	Date range	Digitised item	Item ID	Format
<input type="checkbox"/>	B2455	WEST ERNEST ROBERT	WEST Ernest Robert : Service Number - 5091 : Place of Birth - Benambra VIC : Place of Enlistment - Sale VIC : Next of Kin - (Mother) WEST Mrs F M <a href="#">Access status: Open</a> <a href="#">Location: Canberra</a>	1914 - 1920	 <a href="#">11 pages</a>	8380875	
<input type="checkbox"/>	B2455	WEST ARTHUR	WEST Arthur : Service Number - 7342 7432 : Place of Birth - Sydney NSW : Place of Enlistment - Liverpool NSW : Next of Kin - (Wife) WEST Katie <a href="#">Access status: Open</a> <a href="#">Location: Canberra</a>	1914 - 1920	 <a href="#">56 pages</a>	8380892	
<input type="checkbox"/>	B2455	WEST CHARLES	WEST Charles : Service Number - Depot : Place of Birth - Ballina NSW : Place of Enlistment - Brisbane QLD : Next of Kin - (N/A) BEAVER Ida <a href="#">Access status: Open</a> <a href="#">Location: Canberra</a>	1914 - 1920	 <a href="#">8 pages</a>	8380907	
<input type="checkbox"/>	B2455	WEST FREDERICK GEORGE	WEST Frederick George : Service Number - 4520 : Place of Birth - North Fitzroy VIC : Place of Enlistment - Melbourne VIC : Next of Kin - (Wife) WEST Lillian F <a href="#">Access status: Open</a> <a href="#">Location: Canberra</a>	1914 - 1920	 <a href="#">22 pages</a>	8380922	
<input type="checkbox"/>	B2455	WEST FREDERICK HORACE	WEST Frederick Horace : Service Number - 5769 : Place of Birth - Charters Towers QLD : Place of Enlistment - Charters Towers QLD : Next of Kin - (Mother) WEST Kate <a href="#">Access status: Open</a> <a href="#">Location: Canberra</a>	1914 - 1920	 <a href="#">17 pages</a>	8380923	

<https://recordsearch.naa.gov.au/scripts/AutoSearch.asp?Number=B2455>

WEST Ernest Robert : Service Number - 5091 : Place of Birth - Benambra VIC : Place of Enlistment - Sale VIC : Next of Kin - (Mother) WEST Mrs F M

```
1 Name: WEST Ernest Robert
2 Service Number: 5091
3 Place of Birth: Benambra VIC
4 Place of Enlistment: Sale VIC
5 Next of Kin: (Mother) WEST Mrs F M
```

structured data!



# Mapping our Anzacs



Find a service person and  
see their service record.



Add a note or photograph  
to the scrapbook.

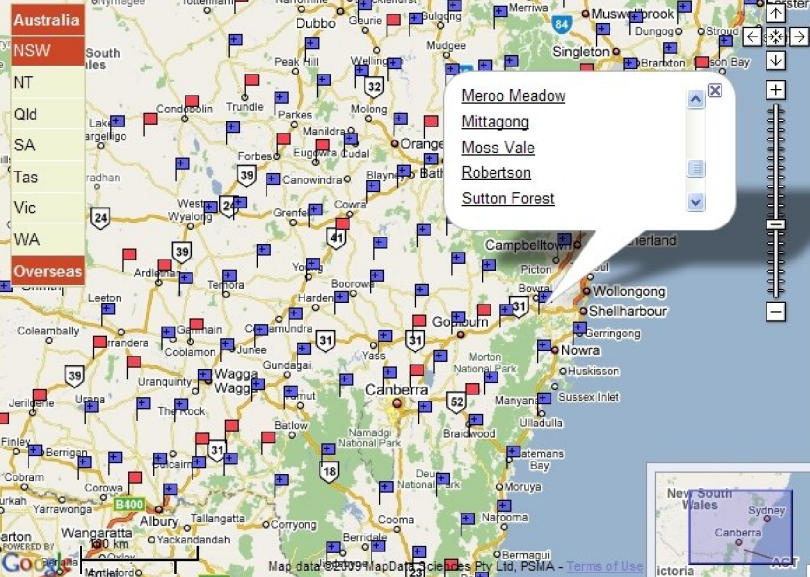


Build a tribute to people  
important to you.

You're not logged in.

[Log in >](#)

[Register >](#)



Click on a **red flag**  
to find people from  
that place.

A **blue flag** marks a  
cluster of places.  
Click for a list of  
places clustered at  
that point.

See what the  
numbers mean.

Or see the other  
search options.

[About this site](#) | [Contact us](#) | [Privacy](#) | Developed with the support of the [Department of Veterans' Affairs](#)

## SHELL-SHOCKED

AUSTRALIA AFTER ARMISTICE





Kevin



me



wragge Merge pull request #3 from wragge/update-readme ··· b19e0ec on 2 Feb 🕒 49 commits

📄 .gitignore	Changes to get access decision reasons	7 years ago
📄 LICENSE.txt	Add licence	7 years ago
📄 README.md	Add link to GLAM Workbench in rs_tools readme	10 months ago
📄 __init__.py	Added init file	9 years ago
📄 client.py	Change Barcode to Item ID	10 months ago
📄 tests.py	Changes for Python 3	4 years ago
📄 utilities.py	Changes for Python 3	4 years ago

☰ README.md ✎

## RecordSearch Tools

This is a Python library for getting useful, machine-readable data out of RecordSearch, the National Archives of Australia's online database.

It's basically a series of screen scrapers and as such is inherently fragile -- changes to the web interface can easily break the code.

RecordSearch is a complex system and I've added bits and pieces to this code over many years. As a result it could do with a major clean-up.

But having said all that, it should work pretty well, and until the NAA provides an API it's really the only way to get data out of RecordSearch.



## RecordSearch Data Scraper

Tools for extracting data from the National Archives of Australia's online database, RecordSearch.

### Table of Contents

[Install](#)  
[How to use](#)

The National Archives of Australia's online database, RecordSearch, contains lots of rich, historical data. Unfortunately there's no API, so we have to resort to screen scrapers to get it out in reusable form. This is a library of scrapers to extract data about the main entities in RecordSearch – Items, Series, and Agencies – from both individual records, and search results.

The main classes are:

- `RSItem()` – an individual item
- `RSItemSearch()` – an advanced search for items
- `RSSeries()` – an individual series
- `RSSeriesSearch()` – an advanced search for series
- `RSAgency()` – an individual agency
- `RSAgencySearch()` – an advanced search for agencies

RecordSearch makes use of an odd assortment of sessions, redirects, and hidden forms, which make scraping a challenge. Please let me know if something isn't working as expected, as problems can be difficult to pin down!

This is a replacement for the original Recordsearch\_tools library. The main changes are:

- Requirements have been updated (dropping RoboBrowser which seems to be no longer maintained)
- The full range of search parameters are now supported for Items, Series, and Agencies
- There's a built-in cache for improved efficiency and speed

See the [documentation](#) for more details. And check out the [RecordSearch](#) section of the GLAM Workbench for examples of what's possible.

[https://github.com/wragge/recordsearch\\_data\\_scraper](https://github.com/wragge/recordsearch_data_scraper)



Book No. 119

C 13/2661

Form No. 21. COMMONWEALTH OF AUSTRALIA.  
DUPLICATE. Immigration Restriction Acts 1901-1905 and Regulations.

No. 042

## CERTIFICATE EXEMPTING FROM DICTATION TEST.

I, William Lawson <sup>actg</sup> the Collector of Customs  
for the State of NEW SOUTH WALES in the said Commonwealth,  
hereby certify that George Say  
hereinafter described, who is leaving the Commonwealth temporarily, will be excepted  
from the provisions of paragraph (a) of Section 3 of the Act if he returns to the Com-  
monwealth within a period of THREE YEARS from this date.

Date 6 May 13 W B Lawson <sup>actg</sup> Collector of Customs.

## DESCRIPTION.

Nationality Chinese Birthplace Canton  
Age 40 years Complexion Dark  
Height 5 ft 8 in Hair Brown  
Build Medium Eyes Brown

Particular marks.

(For impression of hand see back of this document.)

Full Face:—

Profile:—



Date of departure 4 May 13 Destination Fiji  
Ship Levuka  
Date of return 21 July 14 Ship Levuka  
Port Sydney

W B Lawson  
Customs Officer

IMPRESSION OF LEFT HAND.



NAA: ST84/1, 1913/119/41-50



1904/181-190

1904/191-200

1904/21-30

1904/211-220

<http://iabrowse.herokuapp.com/series/ST84-1/browse/>



# THE REAL FACE OF WHITE AUSTRALIA



<https://www.realfaceofwhiteaustralia.net/faces/>

See: Sherratt & Bagnall, 'The people inside' in *Seeing the Past with Computers*

ASKING NEW  
QUESTIONS

## Find an article

Search articles

[Advanced Search](#)

## Find an issue

### by Title

1. The Sydney Morning Herald
2. The Argus
3. Northern Territory Times and Gazette
4. Brisbane Courier
5. The Australian Women's Weekly

Show all titles

### by State

NT QLD  
WA NSW  
SA ACT  
TAS VIC



### by Date

1803

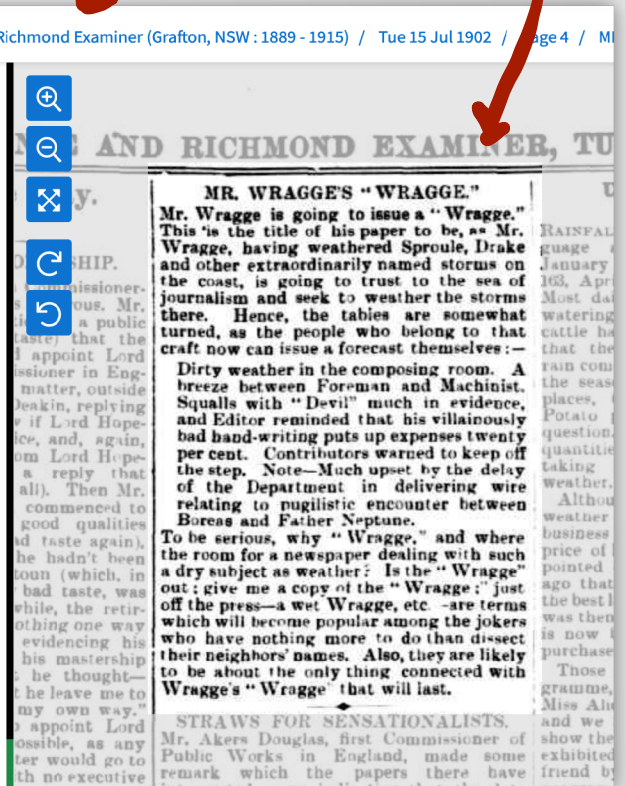
JAN	FEB	MAR	APR
MAY	JUN	JUL	AUG
SEP	OCT	NOV	DEC

metadata

image

OCRd text

MR. WRAGGE'S "WRAGGE."  
Mr. Wragge is going to issue a "Wragge."  
This is the title of his paper to be, as Mr. Wragge, having weathered Sproule, Drake and other extraordinarily named storms on the coast, is going to trust to the sea of journalism and seek to weather the storms there. Hence, the tables are somewhat turned, as the people who belong to that craft now can issue a forecast themselves :- Dirty weather in the composing room. A breeze between Foreman and Machinist. Squalls with "Devil" much in evidence, and Editor reminded that his villainously bad hand-writing puts up expenses twenty per cent. Contributors warned to keep off the step. Note-Much upset by the delay of the Department in delivering wire relating to pugilistic encounter between Boreas and Father Neptune.

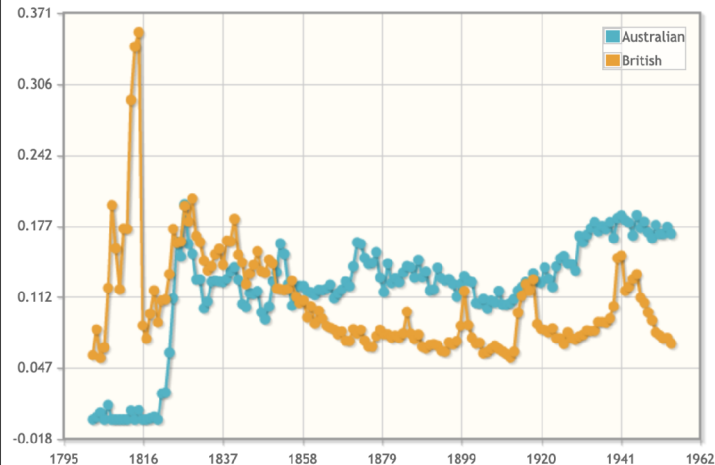




# 2011

## Australian or British

The proportion of Australian newspaper articles including the word 'Australian' versus the proportion including either 'Briton' or 'British' between 1803 and 1954.



- Click and drag to zoom.
- Double click to reset zoom.
- Click on a point to view articles

These calculations are based on the number of articles currently available through [Trove](#). The totals were harvested using my [TroveTools](#) scraper and the lists of articles are generated using my [unofficial Newspapers API](#). Comments or questions to [@wragge](#)

## Results from Trove

Click on a point to see a sample of the articles.

# 2021

## 4. Create your chart

You can add as many queries as you want to a single chart.

Visualise query

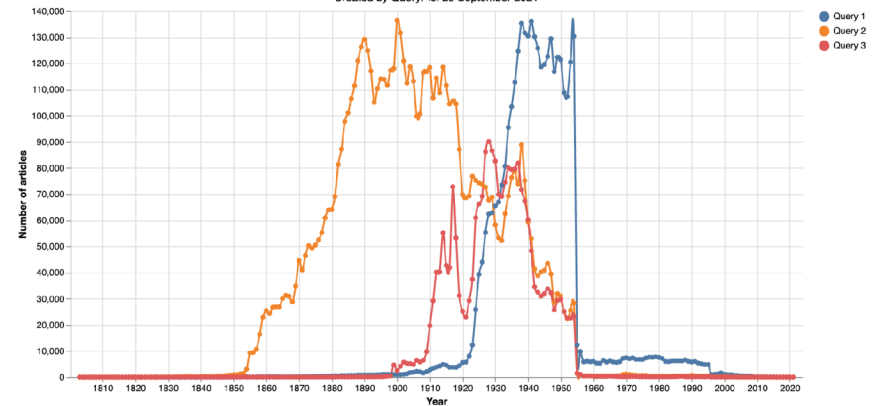
Remove last query

Clear all queries

Raw number of results

### Trove Newspapers & Gazettes Search

Created by QueryPc: 29 September 2021



Query 1 - <https://trove.nla.gov.au/search/category/newspapers?keyword=radio>

Query 2 - <https://trove.nla.gov.au/search/category/newspapers?keyword=telegraph>

Query 3 - <https://trove.nla.gov.au/search/category/newspapers?keyword=wireless>

Save chart as HTML

Width 700

Height 400

Download data: <data/querypc-20210929043718.csv>

<https://glam-workbench.net/trove-newspapers/>

# 2011

## QUERY HARVESTER

We all dip into the Trove Australian newspapers database with a sense of wonder and gratitude. There are such riches to be found. But what if you want to do more than merely dip and pick?

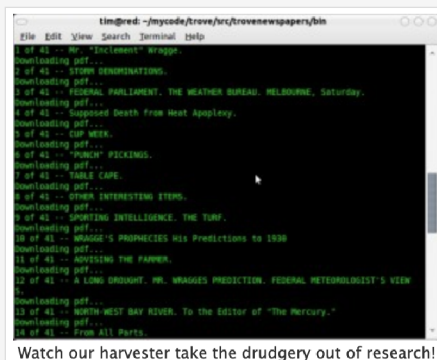
Our query harvester allows you to push beyond the browser, to mine deeper, to extract more. Harvest the details of hundreds, or thousands, or newspaper articles and save them for detailed analysis at your leisure.

Depending on the options you choose, just feed a Trove newspaper query to our harvester and watch in awe as it creates:

- a CSV file containing all the article metadata, ready to be opened as a spreadsheet, or imported into the database of your choice
- a zip file conveniently packaging the text contents of all the harvested articles for analysis program
- a zip file containing pdf versions of all the articles, saved for future reference

Warning: depending on the query, your zip files could get very large!

- [Discover more details](#) (via discontents)
- [Join the discussion](#) (via Trove forum)



# 2021

### Enter your search query

Use the [Trove web interface](#) to construct your search. Remember that the harvester will get **all** of the matched results, not just the first 2,000 you see in the web interface. Once you're happy with your search, just copy the url and paste it below.

Query url:

### Set harvest options

By default the harvester only saves the metadata (date, page, title, newspaper etc) from the search results. If you want to save the full text content of each article, just check the ☐ **Text** box. You can also save PDF copies of every article by checking the ☐ **PDF** option, but be warned that this will slow down your harvest and generate large download files. If you want to save PDFs from large harvests, you're probably better off installing and running the harvester on your own computer.

☐ Save full text

☐ Save PDFs (this can be slow)

Once your harvest is complete a link will appear to download the results as a single, zipped file. See [this notebook](#) for more information about the contents and format of the results folder.

You can also start to explore your results [using this notebook](#).

Created by [Tim Sherratt \(@wragge\)](#) as part of the [GLAM Workbench project](#).

If you think this project is worthwhile you can [support it on Patreon](#).

<https://glam-workbench.net/trove-harvester/>

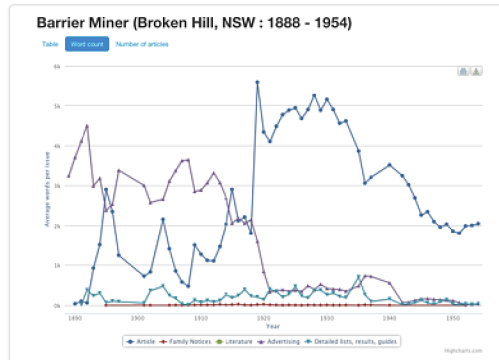


# The front page

The front page of the newspaper is where the 'big' stories are printed. But this hasn't always been the case. The design and function of the front page has changed greatly over time. Here you explore this change...

[Read more at discontents »](#)

## Newspapers

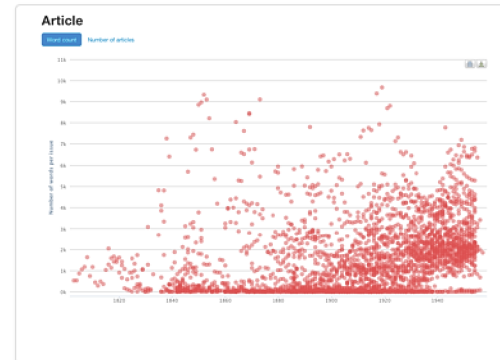


Follow changes in the design of specific newspapers.

Choose a newspaper

Show

## Categories

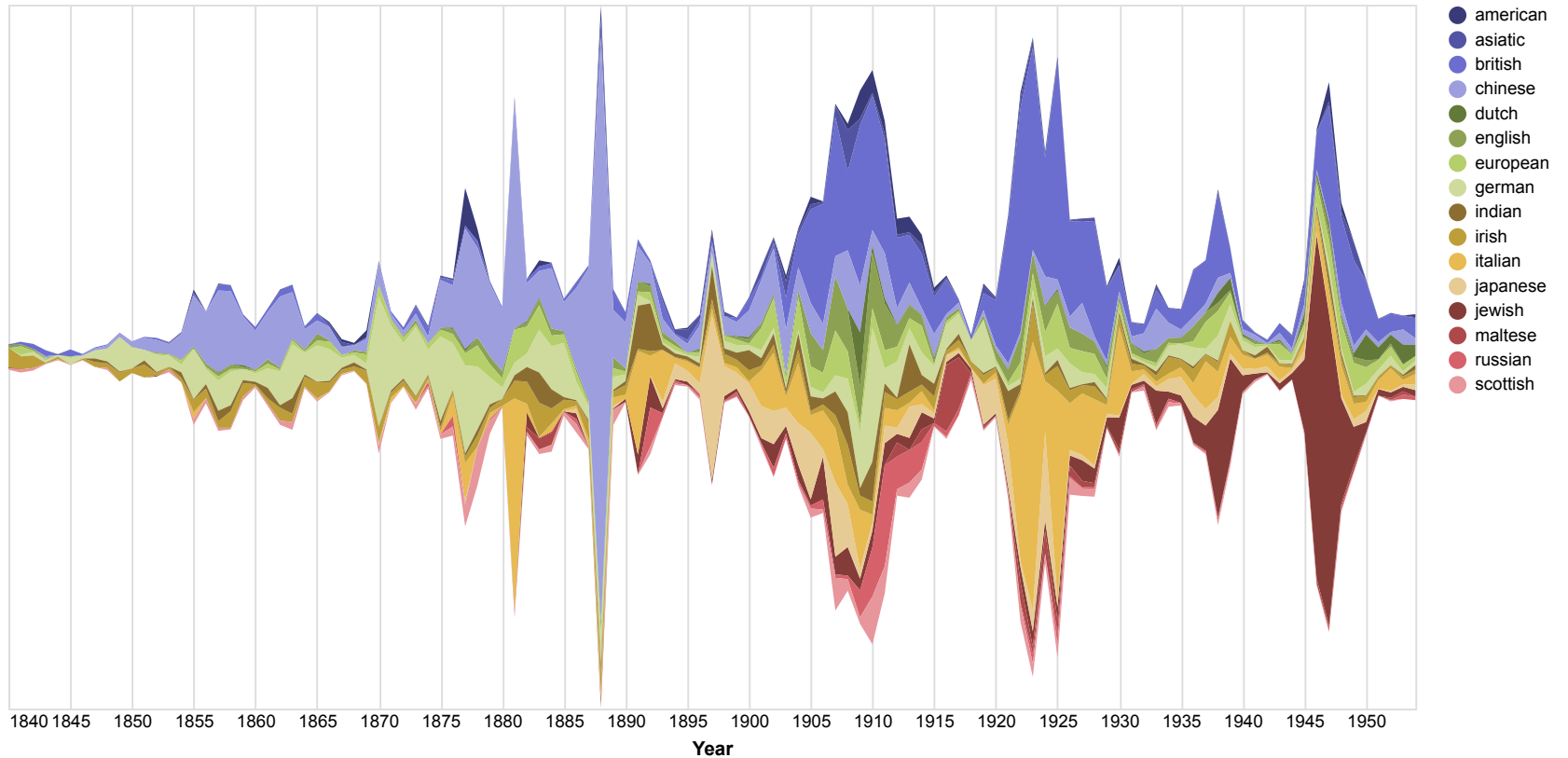


Explore the changing frequency of different types of article.

Choose a category

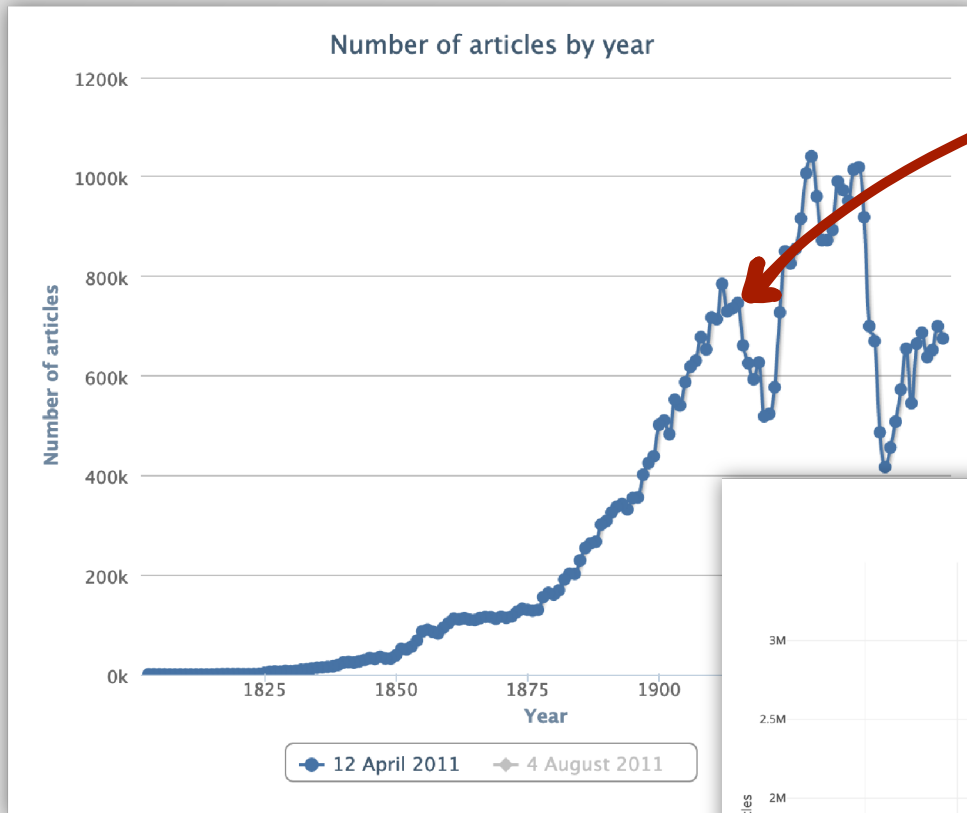
Show

Frequency of words describing origins of immigrants in Trove newspapers

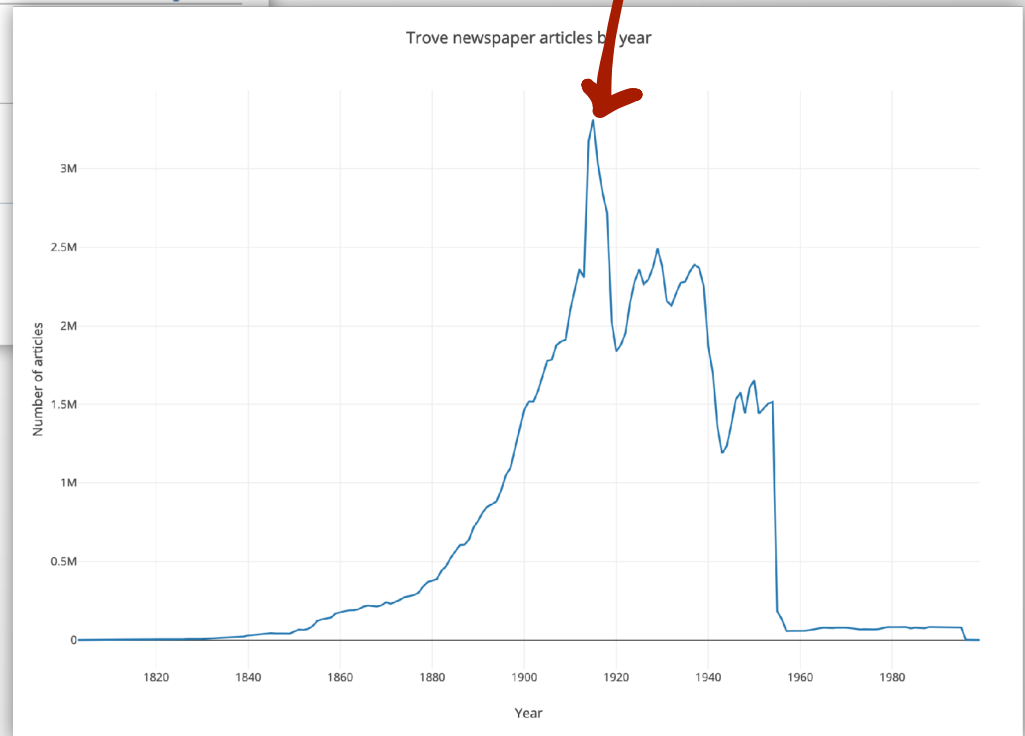


<http://timsherratt.org/blog/who-belongs/>

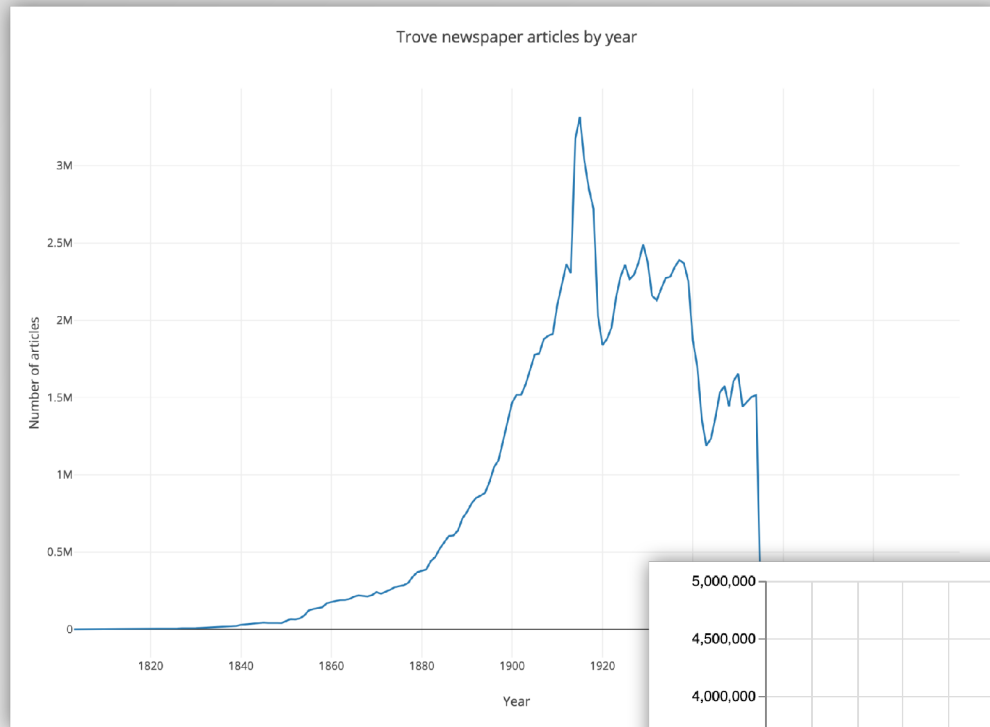
# UNDERSTANDING ACCESS



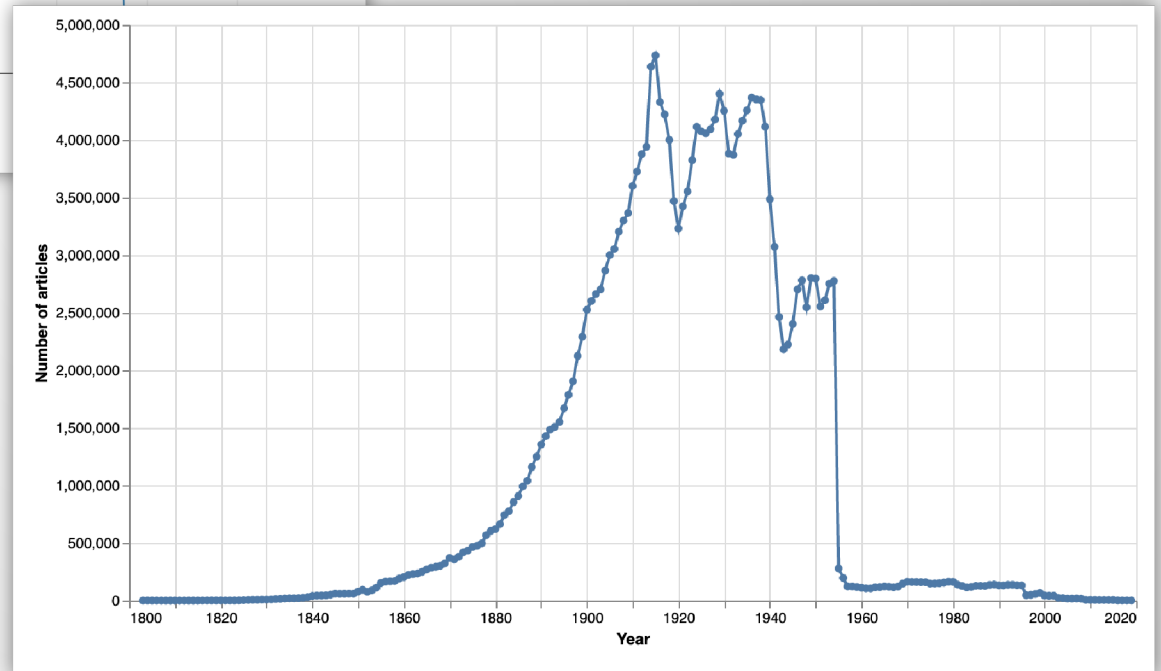
2011



2014



2014



2021



## Overview

Files

Reasons

Ages

Series

Decisions

Harvests

## Closed access

### Overview

Under the Australian *Archives Act 1983* most Commonwealth records are opened to public scrutiny after twenty years (this was reduced from thirty years in 2010). But the Act also defines 'exempt' records that can be withheld from the public for a variety of reasons, including the defence of national security, and the protection of individual privacy. Access under the Act is not an inevitable destination, but a process that may result in records with the access status of 'closed'.

Here you can explore these closed files. Why can't we look at them? How old are they? What are we really being protected against?

14370

**closed files**  
on 1 January 2016

28

**reasons**  
why these files have been closed

58

**years**  
is the average age of these files  
based on their date of their earliest  
content

1128

**series**  
contain closed files

<http://closedaccess.herokuapp.com/>

**INSIDE STORY**

ABOUT SUBSCRIBE LOGIN


**NATIONAL AFFAIRS**

**WITHHELD, PENDING ADVICE**

TIM SHERRATT  
2 FEBRUARY 2018

1492 words

Three snapshots of Australia's national archives reveal delays and anomalies in public access



Greater timeliness on the way? The National Archives of Australia headquarters (currently under renovation) in Canberra. David Coleman/Murray

On the first of January every year, history hits the headlines. That's when the National Archives of Australia, or NAA, releases the most recent batch of cabinet records, documenting discussions at the highest level of government. This year, records from 1994 and 1995 were made available for the first time. Among other things, we learnt that prime minister Paul Keating's government didn't tackle climate change with great gusto.

But what about the files we're not allowed to see? During 2017 at least 1140 government files were withheld from the public. Amid the celebrations each year, we should perhaps take time to investigate the limits of access.

A researcher wanting to see a file from 1967 entitled "Hong Kong — Communist Activities Within the Colony" could have finished two PhDs in the time it took for the file to be processed.

SHARE

**Tim Sherratt**  
@wragge

It's January 1, the day each year when our minds turn to newly released Cabinet records from @naagovau. But while the media focuses on the records that have been made open, I'll be spending the day looking at those that were closed. What weren't you allowed to see in 2020?

11:38 AM · Jan 1, 2021 · TweetDeck

View Tweet activity

155 Retweets 24 Quote Tweets 497 Likes



**Tim Sherratt** @wragge · Jan 1

Replying to @wragge

This will be a "slow" thread, as I gradually pull the data together and document things. But this year I'll be sharing all the data and code through the #GLAMWorkbench, so stay tuned...

2 41

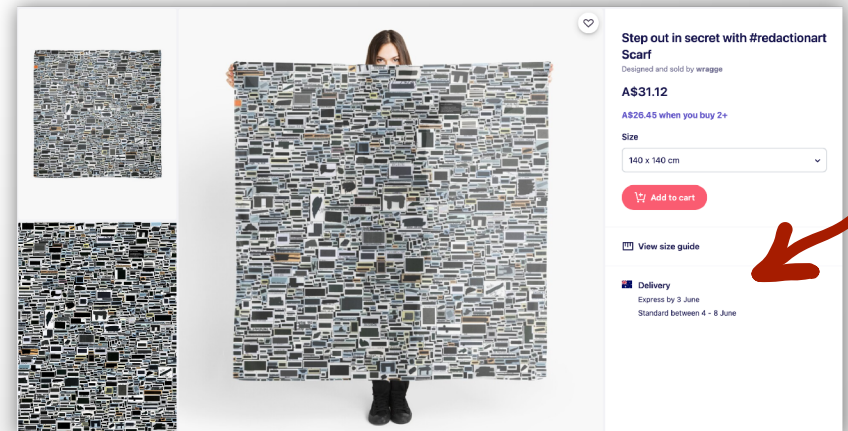
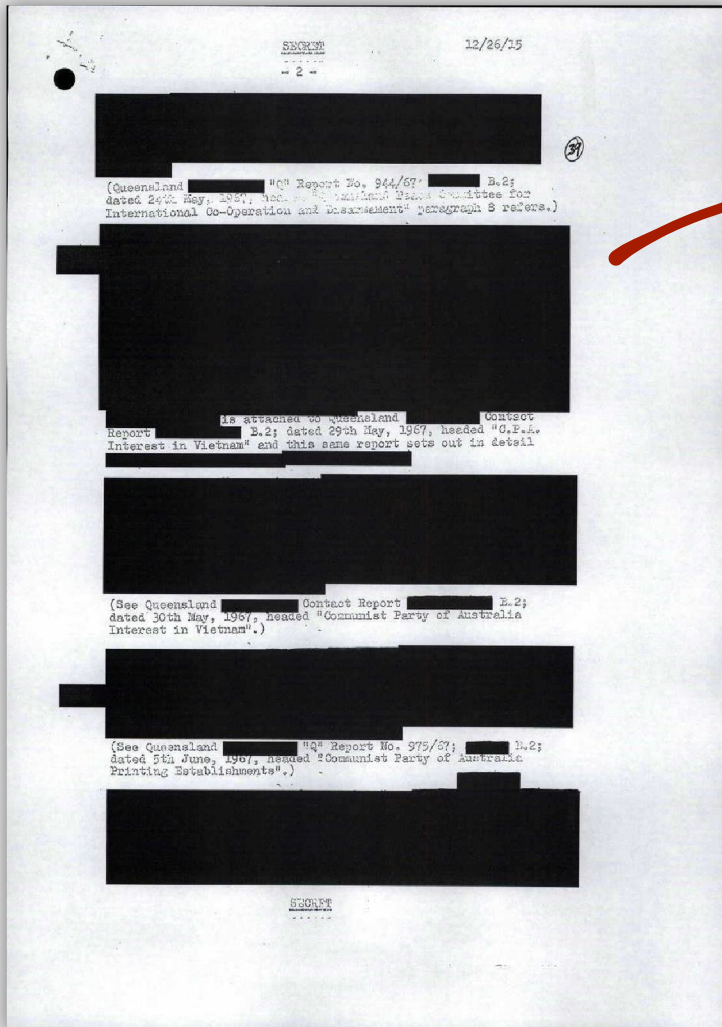
**Tim Sherratt** @wragge · Jan 1

This'll be the sixth consecutive year in which I've harvested all NAA files with an access status of 'closed' on or about 1 January. For some background and past analyses, see my @insidestorymag article from 2018:



**Withheld, pending advice**  
Three snapshots of Australia's national archives reveal delays and anomalies in public access, says Tim Sherratt  
insidestory.org.au

1 11 47



<https://updates.timsherratt.org/2021/04/21/secrets-and-lives.html>



[illegible]

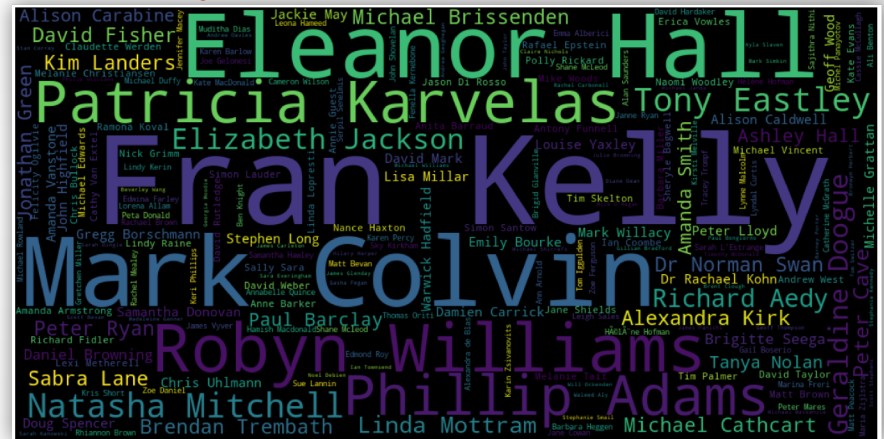
<https://github.com/wragge/diy-redactionart>

# CREATING PATHWAYS

# GLAM DATA IS...

- collection metadata
- text
- images, sound, & video
- 3d models
- user generated
- born digital
- extracted & enriched
- & more...

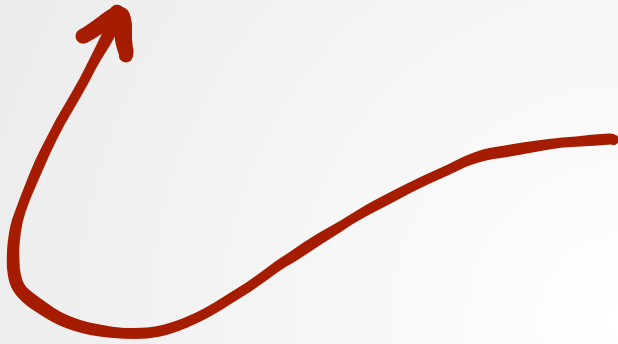
ABC RN  
presenters




<https://glam-workbench.net/trove-music/>





GLAM DATA



RESEARCHERS


GLAM Workbench


Search



[glam-workbench.github.io](https://glam-workbench.github.io)

**GLAM Workbench**  
[Home](#)  
[About](#) >  
[Help](#) >  
[Data sources](#) >  
[Trove](#) >  
[DigitalNZ](#)  
[Archives](#) >  
[Libraries](#) >  
[Web Archives](#)  
[Museums](#) >  
[Government](#) >  
[GLAM Name Index Search](#)  
[Suggest a topic](#)  
[Ask a question](#)  
[Book a chat](#)  
[Latest news!](#)  
[Sponsor](#)

## Welcome to the GLAM Workbench

**Winner of the British Library Labs Research Award 2020!**

Here you'll find a collection of tools, tutorials, examples, and hacks to help you work with data from galleries, libraries, archives, and museums (the GLAM sector). The primary focus is Australia and New Zealand, but new collections are being added all the time. Let me know if there's some GLAM data you'd like me to explore – [suggestions](#) are always welcome!



<https://glam-workbench.github.io/truve-journals/>

**Bulletin cartoons**

- 3,471 images
- 1886 to 1952

15

**Table of contents**  
[Quick start](#)  
[Finding GLAM data](#)  
[Harvesting data](#)  
[Data sources](#)  
[Asking different questions](#)  
[Hacking heritage](#)  
[Bringing documentation alive](#)  
[Do I need to be able to code?](#)  
[Cite as](#)

### Quick start

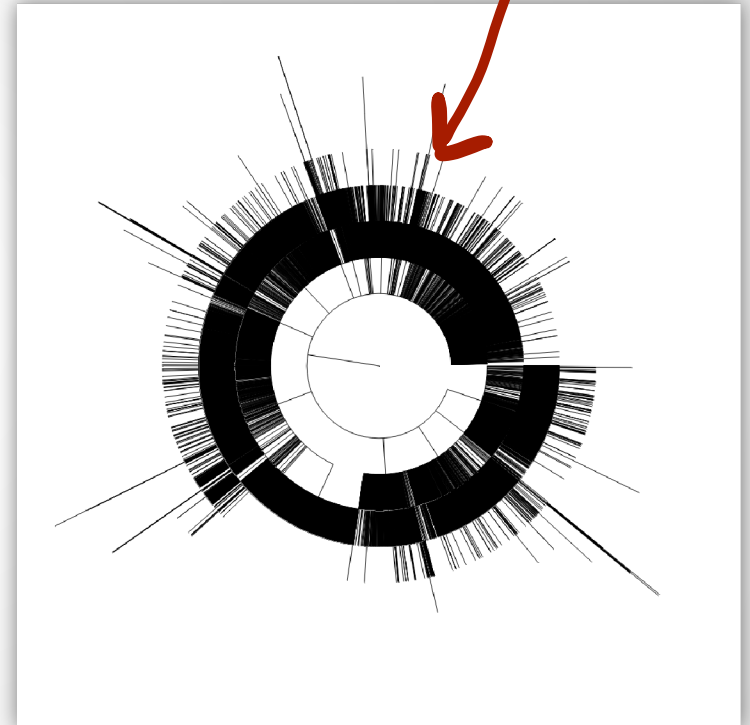
The resources in the GLAM Workbench are created and shared as [Jupyter](#) notebooks. Jupyter lets you combine narrative text and live code in an environment that encourages you to learn and explore. Jupyter notebooks run in your browser, and you can get started without installing any software!

If you want to dive straight in, just have a look around the site and click on one of the links that says 'Run live on Binder'. This will open the notebook, ready to use, in a customised computing environment using the [Binder](#) service.

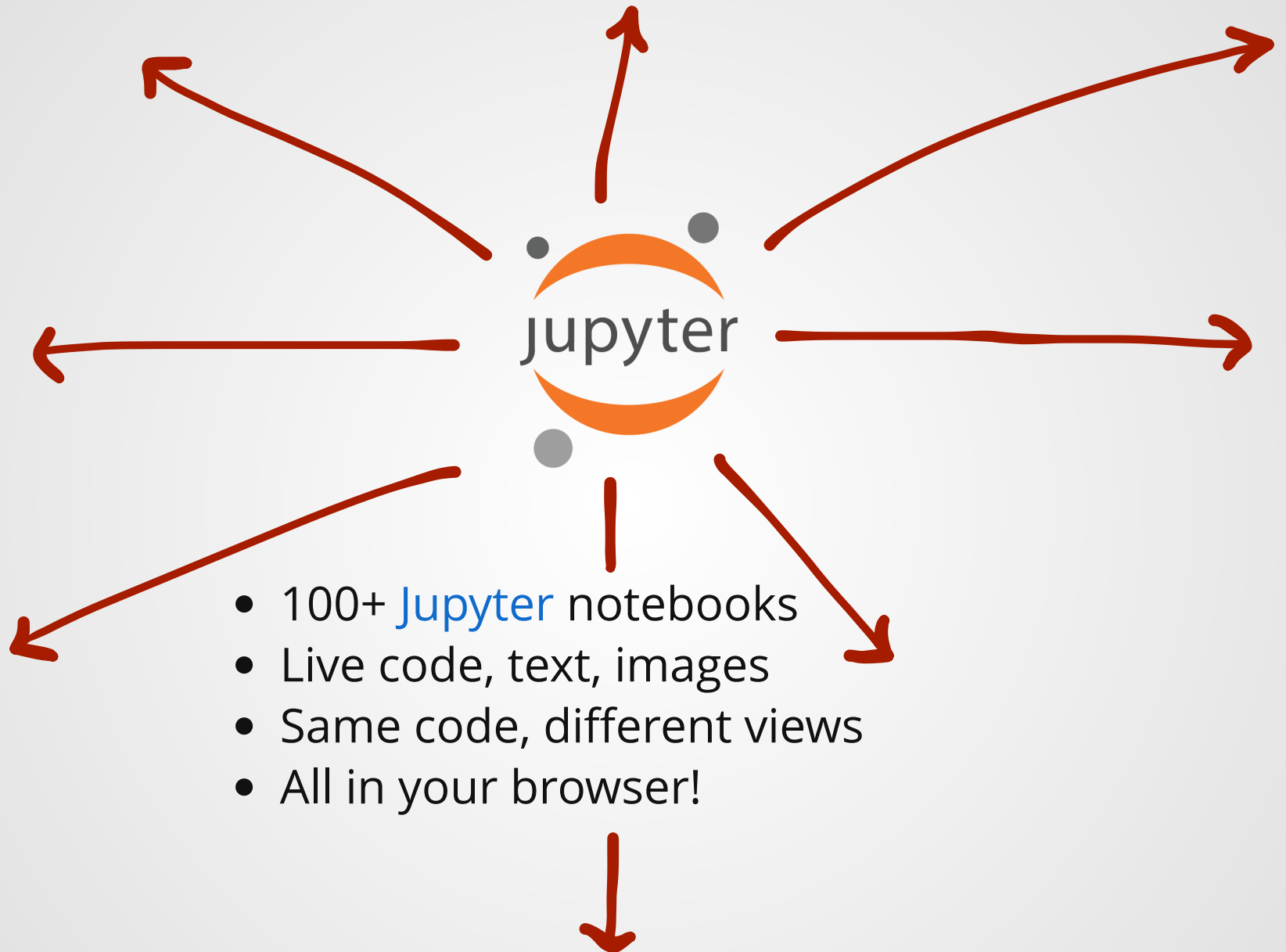
# COLLECTIONS

- Trove
- DigitalNZ
- Commonwealth Hansard
- NAA RecordSearch
- NMA
- Te Papa
- Open data portals
- Web Archives
- & more!

gov.au  
subdomains



<https://glam-workbench.net/web-archives/>



THE HOW

&

THE **WHY**

# JUPYTER NOTEBOOK

code

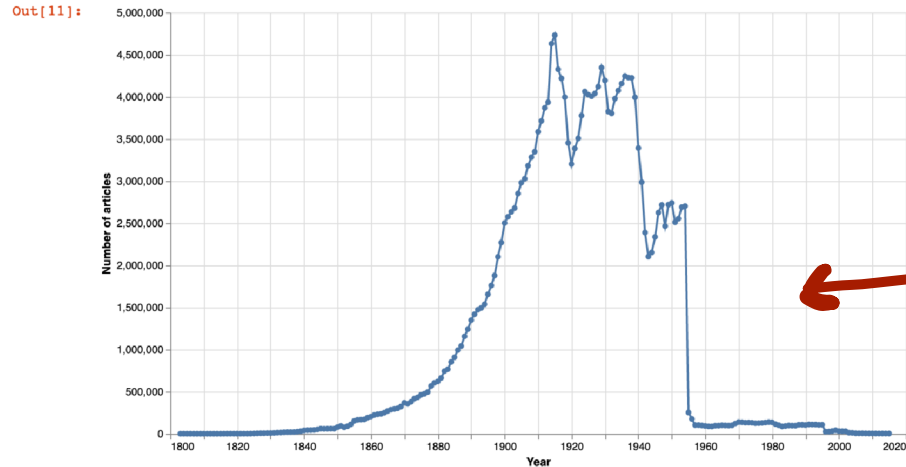
Let's display the results as a simple line chart using Altair

```
In [11]: # Feed Altair our dataframe
alt.Chart(df_total).mark_line(point=True).encode(

    # Years along the X axis
    x=alt.X('term:Q', axis=alt.Axis(format='c', title='Year')),

    # Number of articles on the Y axis (formatted with thousands separators)
    y=alt.Y('total_results:Q', axis=alt.Axis(format=',d', title='Number of articles')),

    # Use tooltips to display the year and number of articles when you hover over a point
    tooltip=[alt.Tooltip('term:Q', title='Year'), alt.Tooltip('total_results:Q', title='Articles', format=',')]
).properties(width=700, height=400)
```



chart

Hmmm, that *is* interesting. There's a significant peak in the number of articles around 1915. Why might that be? Were there more newspapers? Were more articles written because of the war?

Nope. It's because of funding and digitisation priorities. Not all Australian newspapers are in Trove. Some have been lost, and many are just waiting to be digitised. Funding is always limited, so priorities have to be set. In the lead up to the centenary of World War I, it was decided to focus on digitising newspapers from that period. This chart reflects those priorities. This is not the number of newspaper articles published in Australia, it's the number of newspaper articles that have been digitised and made available through Trove. It's important to remain aware of this as you use Trove.

text



# Using Jupyter notebooks in the GLAM Workbench

The GLAM Workbench includes many [Jupyter notebooks](#). Jupyter lets you combine text, images, and live code within a single web page. So not only can you read about collections data, you can download it, analyse it, and visualise it – all within your browser!

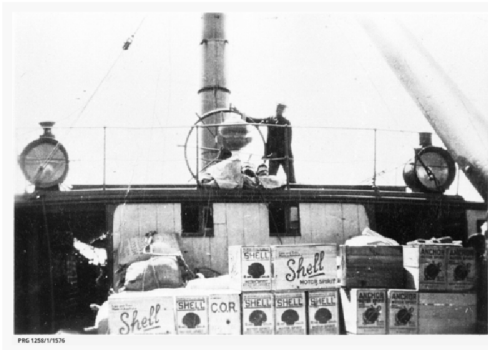
While notebooks often include some fairly intimidating looking code, you don't need to understand the code to use them. As explained below, there's just a couple of basic conventions you need to keep in mind when running Jupyter notebooks. Once you've mastered these, you'll be able to use any of the tools or examples in this workbench.

Of course, once you've developed a bit of confidence, you might want to start playing around with the code. That's how you learn. The GLAM Workbench isn't just a collection of tools, it's a starting point – from here you can explore, extend, and experiment!

## 1. Text cells and code cells

You're currently inside a Jupyter notebook. It might look like any old web page, but there are a number of important differences.

Jupyter notebooks contain two main components – **text cells** and **code cells**. This is a text cell. As well as ordinary text, these cells can contain things like links, images, and tables. Text cells are used to provide documentation and context.



*Text cells can include images and links. All aboard the PS Jupiter, thanks to the [State Library of South Australia](#)!*

**Code cells** are easy to identify, as they have boxes around them, like this:

```
[ ]: # Hi there! I'm a code cell!  
  
# Lines that begin with a hash symbol, like these, are comments within code.  
# They don't do anything, but they help you understand how the code works.
```

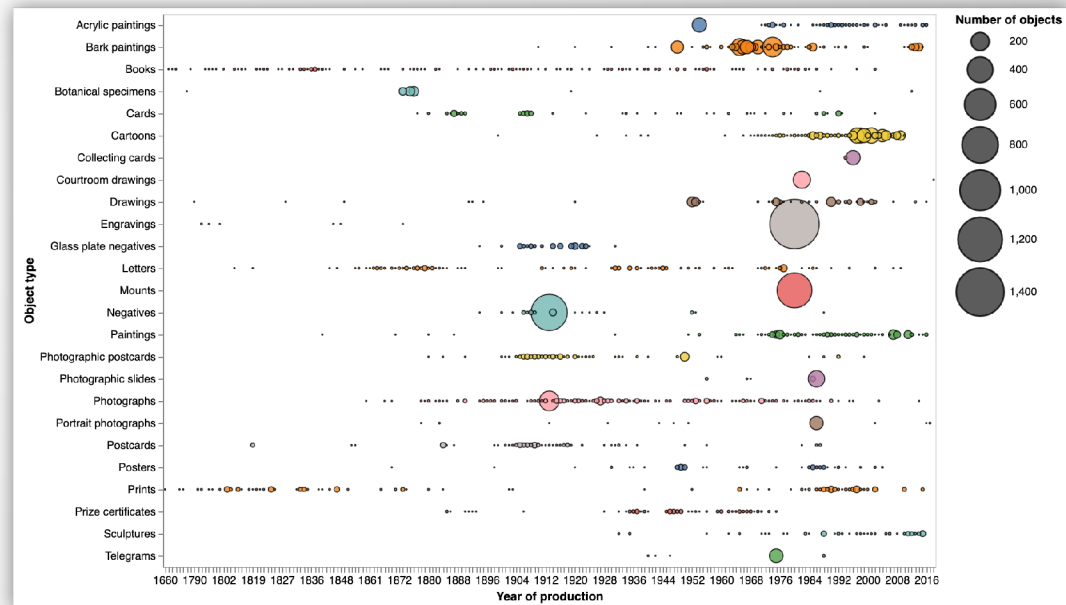
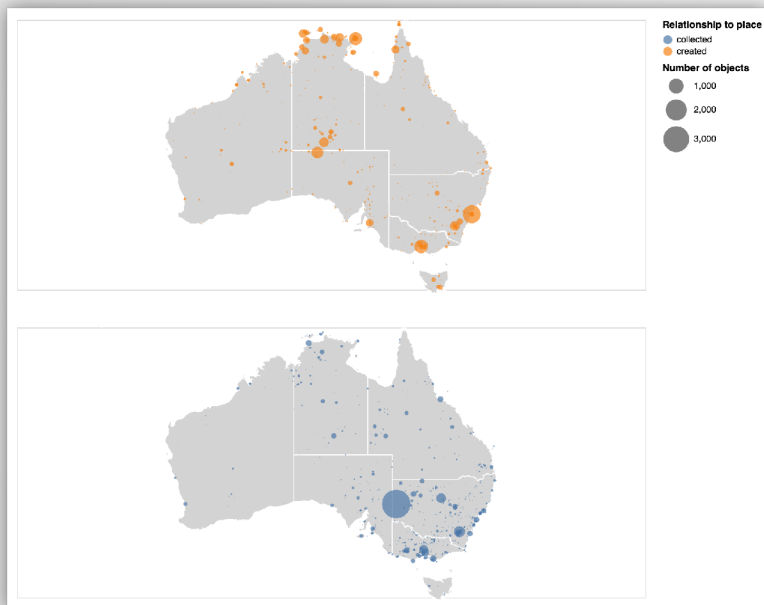
Code cells contain snippets of live, executable code – they **do things**. It's this combination of documentation and real, live code examples that make Jupyter notebooks so useful. You can learn by doing!

# GETTING STARTED

→ TRY IT!

<https://glam-workbench.net/getting-started/>

# NATIONAL MUSEUM OF AUSTRALIA



<https://glam-workbench.net/nma/>

# EXPLORE APIS

- Trove
- DigitalNZ
- National Museum of Australia
- Te Papa
- Museums Victoria
- Web Archives

A random item from the Museum of Victoria

```
[67]: import requests
import random
from IPython.display import Image, display, HTML
import ipywidgets as widgets


[36]: SEARCH_URL = 'https://collections.museums.victoria.com.au/api/search'

[72]: def display_random(b):
    out.clear_output()
    params = {
        'query': '',
        'sort': 'random',
        'hasImages': 'yes'
    }
    response = requests.get(SEARCH_URL, params=params)
    record = random.choice(response.json())
    with out:
        display(HTML('<h3>{record["displayTitle"]}</h3>'))
        display(Image(url=record["media"][0]["small"]["uri"]))
        display(HTML('<a href="https://collections.museums.victoria.com.au/{record["id"]}">More info...</a>'))

    go = widgets.Button(description="Randomise!")
    out = widgets.Output()
    go.on_click(display_random)
    display(go)
    display(out)
```

Randomise!

Proof Coin - 1/2 Cent, British Caribbean Territories, 1955



[More info...](#)

# GLAM Jupyter resources

Here are some more Jupyter notebooks and resources for exploring GLAM data.

## GLAM insitutions

- [ACMI public API – Interactive examples](#)
- [Data exploration](#), Library of Congress Labs
- [Data Foundry](#), National Library of Scotland
- [Datasets and Digital objects](#), Deutsche National Bibliotek
- [GLAM Jupyter Notebooks](#), Biblioteca Virtual Miguel de Cervantes Labs
- [Jupyter Notebooks using the British Library's Digital Collections & Data](#), British Library
- [V&A Collections API Guide](#) and [Data Explorations](#), Victoria and Albert Museum



### Table of contents

GLAM insitutions  
Teaching and learning  
Other lists  
Contributing to this list

<https://glam-workbench.net/more-glam-notebooks/>

PLACES TO START

&

PLACES TO GO

[View code at GitHub](#) · [View details in GLAM Workbench](#)

This notebook provides a quick overview of CSV-formatted data files, particularly those created by GLAM institutions (galleries, libraries, archives, and museums). The list of CSV files from Australian GLAM insitutions provided below [was harvested from state and national government data portals](#). You can select a file from the list or upload your own.

The CSV Explorer looks at each column in the selected CSV file and tries to identify the type of data inside. It then attempts to tell you something useful about it. There's some more details about the process below.

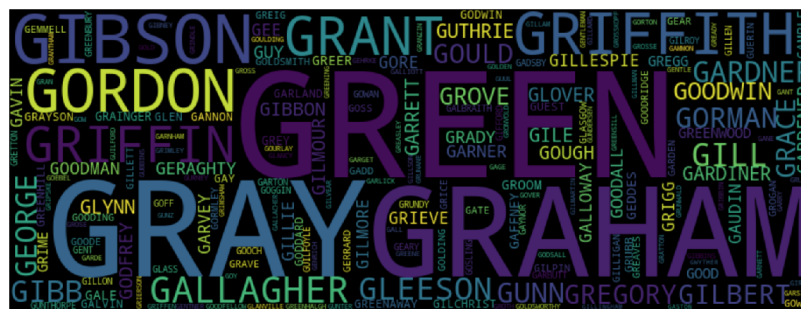
Given all the possible variations in recording and formatting data, there will be errors. But hopefully this will provide you with a useful starting point for further exploration.

1. Last name

## Summary

Number of (non empty) values	12,884 (100.00% of rows)
Number of unique values	1,864 (14.47% of non-empty values)
Total number of words	12891
Highest number of words	2
Median number of words	1.000000
Number of empty records	0

This look like it contains text.





## QueryPic

### Visualise searches in Trove's newspapers and gazettes

What does it mean when your search in [Trove's digitised newspapers](#) returns 3 million results? QueryPic helps you explore your search results by showing you how they change over time.

- Just search in Trove, then copy and paste the url below to visualise the results.
- Combine multiple searches to compare changes in language, technology, or the impact of particular events.
- Click on any point in a chart to view the results for that date in Trove.
- Save your chart as an image or a HTML file.

### 1. Enter your Trove API key

Get your own [Trove API key](#) and enter it below.

API key:

### 2. Enter your search url

Construct your search using the [Trove web interface](#), then just copy and paste the url into the box below.

Query:

### 3. Select the scale (optional)

QueryPic aggregates search results by time units – either 'year', 'month', or 'day'. If you choose 'Automatic' in the list below, QueryPic will choose a unit based on the date range of your query, trying to balance resolution and efficiency:

- Less than 3 months – 'day'
- Between 3 months and 20 years – 'month'
- More than 20 years – 'year'

If you're not happy with these results you can select your own time unit.

Time unit:

### 4. Create your chart

You can add as many queries as you want to a single chart.

[Visualise query](#)

[Remove last query](#)

[Clear all queries](#)

Created by [Tim Sherratt](#) for the [GLAM Workbench](#). Support this project by becoming a [GitHub sponsor](#).

notebook

### 2. Find the number of articles per year using facets

Decade
1950-1959 (3,010)
1940-1949 (5,925)
1930-1939 (12,181)
1920-1929 (11,828)
1910-1919 (15,348)
1900-1909 (29,350)
1890-1899 (34,867)
1880-1889 (7,770)
1870-1879 (3,487)
1860-1869 (5,689)
<a href="#">more...</a>

When you search for newspaper articles using Trove's web interface, there will appear alongside a column headed 'Refine your results'. This column displays summary data extracted from your search, such as the states in which articles were published and the newspapers that published them. In the web interface, you can use this data to filter your results, but using the API we can retrieve the raw data and use it to visualise the complete result set.

Here you can see the decade facet, showing the number of newspaper articles published each decade. If you click on a decade, the interface displays the number of results per year. So sitting underneath the web interface is data that breaks down our search results by year. Let's use this data to visualise a search over time.

To get results by year from the Trove API, you need to set the `facet` parameter to `year`. However, this only works if you've also selected a specific decade using the `1-decade` parameter. In other words, you can only get one decade's worth of results at a time. To assemble the complete dataset, you need to loop through all the decades, requesting the `year` data for each decade in turn.

Let's start with some basic parameters for our search.

```
In [3]: # Basic parameters for Trove API
params = {
    'facet': 'year', # Get the data aggregated by year.
    'zone': 'newspaper',
    'key': 'api_key',
    'encoding': 'json',
    'n': 0 # We don't need any records, just the facets!
}
```

But what are we searching for? We need to supply a `q` parameter that includes our search terms. We can use pretty much anything that works in the Trove simple search box. This includes boolean operators, phrase searches, and proximity modifiers. But let's start with something simple. Feel free to modify the `q` value in the cell below.

```
In [4]: # CHANGE THIS TO SEARCH FOR SOMETHING ELSE!
params['q'] = 'radio'
```

Let's define a couple of handy functions for getting facet data from the Trove API.

<https://glam-workbench.net/trove-newspapers/>

ZOOM OUT

&

ZOOM IN

# TROVE NEWSPAPER HARVESTER

## Enter your search query

Use the [Trove web interface](#) to construct your search. Remember that the harvester will get all of the matched results, not just the first 2,000 you see in the web interface. Once you're happy with your search, just copy the url and paste it below.

Query url:

## Set harvest options

By default the harvester only saves the metadata (date, page, title, newspaper etc) from the search results. If you want to save the full text content of each article, just check the ☐ **text** box. You can also save PDF copies of every article by checking the ☐ **pdf** option, but be warned that this will slow down your harvest and generate large download files. If you want to save PDFs from large harvests, you're probably better off installing and running the harvester on your own computer.

☐ Save full text

☐ Save PDFs (this can be slow)

Start harvest

Once your harvest is complete a link will appear to download the results as a single, zipped file. See [this notebook](#) for more information about the contents and format of the results folder.

You can also start to explore your results [using this notebook](#).

Created by [Tim Sherratt](#) (@wragga) as part of the [GLAM Workbench](#) project.

If you think this project is worthwhile you can [support it on Patreon](#).

```
trove-newspaper-harvester > notebooks > data > 1625733001 > {} metadata.json > ...
1  {
2    "query": "https://trove.nla.gov.au/search/category/newspapers?keyword=s22clement%20wragga%22&1-illustrationType=Photo",
3    "key": "s22clement%20wragga",
4    "max": null,
5    "pdf": false,
6    "text": true,
7    "image": true,
8    "harvest": "1625733001",
9    "date_started": "2021-07-08T18:30:01.156879",
10   "start": null
11 }
```

article_id	title	newspaper_id	newspaper_title	page	date	category	words	illus	coi
10197292	Still a good life, at 107	11	The Canberra Times (ACT : 1926 - 1995)	1	1988-02-28	Article	830	Y	10
10341118	BOWLS	528	The Arrow (Sydney, NSW : 1896 - 1912)	4	1910-03-05	Article	3366	Y	0
104165789	THE INQUIRY DEPARTMENT This Inquiry Department service is free to subscribers, who	470	The Land (Sydney, NSW : 1911 - 1954)	22	1938-05-27	Article	2335	Y	0
104178640	Long-range Weatherman	470	The Land (Sydney, NSW : 1911 - 1954)	96	1937-07-14	Article	291	Y	0
104180732	Irigo Jones Attempts Riddle of the Seasons to Lift the Veil on the	470	The Land (Sydney, NSW : 1911 - 1954)	34	1939-08-23	Article	1514	Y	0
10461344	The Talk of Sydney House Decoration in the Trenches.	438	The Mirror of Australia (Sydney, NSW : 1915 - 191)	7	1915-06-19	Article	1650	Y	1
104905145	All About People Title Table.	447	The Catholic Press (Sydney, NSW : 1896 - 1942)	14	1905-10-19	Article	6033	Y	0
106257749	Recollections, AND The Road of L'esunua.	494	Freeman's Journal (Sydney, NSW : 1850 - 1932)	12	1903-12-19	Article	4529	Y	6
110632120	GANG GANG Everybody wanted to help	11	The Canberra Times (ACT : 1926 - 1995)	3	1979-01-09	Article	1540	Y	0
111093069	PERSONAL	494	Freeman's Journal (Sydney, NSW : 1850 - 1932)	17	1911-09-07	Article	1774	Y	1
11374054	MAN CANNOT YET CAUSE RAIN But Australia Has Seen Some Interesting Experiments	13	The Argus (Melbourne, Vic. : 1848 - 1957)	2	1844-12-09	Article	1113	Y	0
115733964	FORM OF HORSES FOR WARWICK FARM TO-DAY Tukeroo Discusses Programme For	439	Arrow (Sydney, NSW : 1916 - 1933)	1	1917-02-03	Article	4036	Y	0
116730908	ON THE ROOF OF AUSTRALIA: THE NEW PLAYGROUND FOR TIRED PEOPLE. The Sto	464	Freeman's Journal (Sydney, NSW : 1850 - 1932)	3	1917-12-06	Article	2997	Y	1
118782000	TO-DAY'S DIARY OF A MAN ABOUT TOWN	508	Evening News (Sydney, NSW : 1860 - 1931)	4	1923-12-15	Article	571	Y	0
119430205	BILL OF THE PLAY. The Week's Programme. Commencing Saturday.	437	The Newlander: an Australian Paper for Australian	7	1904-02-06	Article	2710	Y	0
119437316	BILL OF THE PLAY. The Week's Programme. Commencing Saturday.	437	The Newlander: an Australian Paper for Australian	7	1904-03-19	Article	2675	Y	0
120096374	SOLUTION OF No. 837.	505	Sunday Times (Sydney, NSW : 1899 - 1930)	21	1911-09-03	Detailed list	163	Y	0
122904091	MIDWEEK MAGAZINE Beauty and danger of mountain newsmaking	11	The Canberra Times (ACT : 1926 - 1995)	23	1991-11-13	Article	1619	Y	0
12520267	WE MAY SOON FORETELL THAT A Queenslandeer is Calculating the Influence of Planets	13	The Argus (Melbourne, Vic. : 1848 - 1957)	8	1938-10-29	Article	2789	Y	2
12616871	PERSONAL	505	Sunday Times (Sydney, NSW : 1899 - 1930)	7	1898-01-08	Article	1236	Y	0
126748804	The Land of the Utmost South	505	Sunday Times (Sydney, NSW : 1899 - 1930)	1	1908-12-13	Article	2667	Y	0
128449339	THE WEEK.	466	The World's News (Sydney, NSW : 1901 - 1955)	16	1902-05-10	Article	2383	Y	2
128796103	Stealing the Public Pools.	437	The Newlander: an Australian Paper for Australian	18	1903-08-22	Article	529	Y	0



Name	Date Modified	Size	Kind	
18840601-1146-224815906.txt	8 Jul 2021 at 6:41 pm	59 bytes	Plain Text	
18900027-804-6361584.txt	8 Jul 2021 at 6:45 pm	14 KB	Plain Text	
18900327-804-63615627.txt	8 Jul 2021 at 6:45 pm	14 KB	Plain Text	
18901122-804-6361877.txt	8 Jul 2021 at 6:45 pm	13 KB	Plain Text	
18901122-804-63615892.txt	8 Jul 2021 at 6:45 pm	13 KB	Plain Text	
18901328-804-63612334.txt	8 Jul 2021 at 6:45 pm	11 KB	Plain Text	
18901328-804-63616375.txt	8 Jul 2021 at 6:45 pm	8 KB	Plain Text	
18902016-42-19820349.txt	8 Jul 2021 at 6:37 pm	15 KB	Plain Text	
18902016-804-64030157.txt	8 Jul 2021 at 6:45 pm	16 KB	Plain Text	
18902083-804-64030125.txt	8 Jul 2021 at 6:46 pm	13 KB	Plain Text	
18940526-1176-22726300.txt	8 Jul 2021 at 6:41 pm	23 KB	Plain Text	
18940818-821-16190812.txt	8 Jul 2021 at 6:33 pm	20 KB	Plain Text	
18940908-698-16263080.txt	8 Jul 2021 at 6:34 pm	14 KB	Plain Text	
18956720-847-198587833.txt	8 Jul 2021 at 6:37 pm	2 KB	Plain Text	
18956803-42-21634432.txt	8 Jul 2021 at 6:40 pm	4 KB	Plain Text	
18970626-1435-25095178.txt	8 Jul 2021 at 6:42 pm	36 KB	Plain Text	
18971127-191-70932349.txt	8 Jul 2021 at 6:46 pm	11 KB	Plain Text	
18971211-319-197307937.txt	8 Jul 2021 at 6:37 pm	5 KB	Plain Text	
18971218-186-68183641.txt	8 Jul 2021 at 6:46 pm	40 KB	Plain Text	
18971225-698-163799609.txt	8 Jul 2021 at 6:34 pm	19 KB	Plain Text	
18980109-005-12518571.txt	8 Jul 2021 at 6:31 pm	7 KB	Plain Text	
18980116-82-7182235.txt	8 Jul 2021 at 6:46 pm	26 KB	Plain Text	
18980206-1161-21189921.txt	8 Jul 2021 at 6:38 pm	12 KB	Plain Text	
18980702-1161-21145085.txt	8 Jul 2021 at 6:39 pm	7 KB	Plain Text	
18980924-42-20861726.txt	8 Jul 2021 at 6:38 pm	2 KB	Plain Text	
18981217-319-197534850.txt	8 Jul 2021 at 6:37 pm	4 KB	Plain Text	
18990316-874-180223434.txt	8 Jul 2021 at 6:36 pm	28 KB	Plain Text	
18990101-647-155451382.txt	8 Jul 2021 at 6:32 pm	6 KB	Plain Text	
19000023-1161-211455556.txt	8 Jul 2021 at 6:39 pm	13 KB	Plain Text	
19000510-713-145935783.txt	8 Jul 2021 at 6:32 pm	9 KB	Plain Text	
19001031-319-145566650.txt	8 Jul 2021 at 6:33 pm	13 KB	Plain Text	

THE PROMENADE. A coat and skirt for ladies are associated in the left hand figure. A handsome suit resembling a strict tailor finish is here shown in a development of side raced cloth. The coat has lapelled fronts, luv, close to a fly, but visible outwear may be as a jilthead, keys are formed at the top of the fronts, and meet the ends of the rolling collar. The basic is closely fitted, and the plait are arranged at the lower part of the side-back seams, flight gathered fullness appears at the top of the sleeves and lap, cover 3d. pockets are inserted in the fronts. A pored circular flounce characterises the skirt, which is of the five pored variety with a habit back. A closing is arranged at the left side front seam, and the mode is adaptable for sweep, round or instep length, homespun, suiting, serge, chevise and all varieties of cloth and- recommended for developing the suit, and stitched strappings of taffeta or satin will give all the trimming necessary. Light tan-faced cloth will develop most isily for Autumn wear, and self colored pan as, or pastel de tete may be used for facing the collar and revers. In the other figure a ladies' skirt and Norfolk jacket are united. Habit coats are prominent among the new skirt designs, and like so mode formic a part of this toilette is unusually attractive. Mixed gray cloth was used for the development, released by velvet strappings over laying plain cloth. Mice skirt com prises five gir3r, and the spring at this lower part of the seams j-r quite pronounced, giving a pretty flare. It is adaptable for L-wesp or rr and length, and a closing may be made at the side or at the back. Norfolk styles are attracting much attention this season, their chic ap pearance and heart air giving them a style all their own. A jaunty design is here pictured produced in rough goods matching the skirt, with vel vet strappings and as is lay of the same on the collar and lapels. Fanci ful yokes are a modish feature, as in below the yokes plain-) are applied toth back and front. A closing is er fected with buttons and button-holes and at the top lapels are formed. A rolling collar is introduced, and a belt passing under the plait closes in front. Strap cuffs complete the sleeves. Homespun in an available f cloth-and white check will make up stylishly. Both velvet facings and black bone bat tons for dosing. Colort cloth, serge, broadcloth, Venetian and the many beaves of faced cloth are recommended. A leather belt lay for use instead of the material. The linen getters are 88 in the 3rd and 1/2th instant were fined 25 each, for getting tin ber

19020428-269-82349152.txt  
Plain Text Document - 12 KB

<https://glam-workbench.net/trove-harvester/>

# HARVESTING DATA

- OCRd text from Trove [books](#) & [journals](#)
- [Front covers of the Women's Weekly](#)
- [Newspapers from Papers Past](#)
- [Press releases from Parliament](#)
- [Tags applied by Trove users](#)
- [Radio National programs](#)
- [Text from archived web pages](#)





# READY-TO-GO DATA!

- 3,471 Bulletin cartoons
- Press releases about refugees
- Segments broadcast on RN Breakfast
- Commonwealth Hansard



# WEB ARCHIVES

## Display changes in the text of an archived web page over time

This notebook displays changes in the text content of a web page over time. It retrieves a list of available captures from a Memento Timemap, then compares each capture with its predecessor, displaying changes side-by-side.

By default, the notebook only displays lines that have changed. If you want to see more context, you can adjust the parameters in the `show_all_differences()` function to show lines around each change, or the complete text content.

Archive:

URL:

[Show text changes](#)

[Clear all](#)

### Key

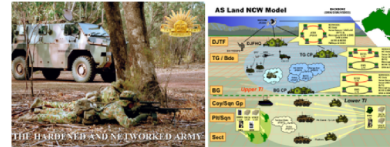
- **deleted text**
- **changed text**
- **added text**

		<a href="#">21 March 2020</a>			<a href="#">22 March 2020</a>
1		No Differences Found	1		No Differences Found

		<a href="#">22 March 2020</a>			<a href="#">22 March 2020</a>
1		No Differences Found	1		No Differences Found

		<a href="#">22 March 2020</a>			<a href="#">23 March 2020</a>
n			n	16	Announcements
				17	Schools and childcare are only open to children of key workers
				18	Pub, restaurants, and leisure venues must stay closed or they may be fined
				19	Food shops and pharmacies will remain open
n			n	27	How to protect extremely vulnerable people (shielding)
n			n	34	School closures, education, and childcare
				35	What parents and carers need to know about closures

## Find and explore Powerpoint presentations from a specific domain



[New to Jupyter notebooks? Try Using Jupyter notebooks for a quick introduction.](#)

Web archives don't just contain HTML pages! Using the `filter` parameter in CDX queries we can limit our results to particular types of files, for example Powerpoint presentations.

This notebook helps you find, download, and explore all the presentation files captured from a particular domain, like `defence.gov.au`. It uses the Internet Archive by default, as their CDX API allows domain level queries and pagination, however, you could try using the UKWA or the National Library of Australia (prefix queries only).

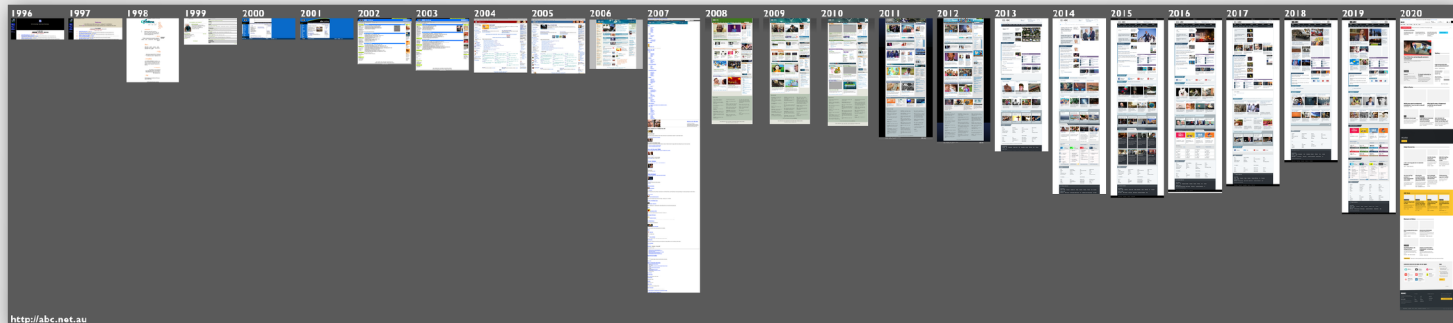
This notebook includes a series of processing steps:

1. Harvest capture data
2. Remove duplicates from capture data and download files
3. Convert Powerpoint files to PDFs
4. Extract screenshots and text from the PDFs
5. Save metadata, screenshots, and text into an SQLite database for exploration
6. Open the SQLite db in Datasheet for exploration

Here's an [example of the SQLite database](#) created by harvesting Powerpoint files from the `defence.gov.au` domain, running in Datasheet on Glitch.

Moving large files around and extracting useful data from proprietary formats is not a straightforward process. While this notebook has been tested and will work running on Binder, you'll probably want to shift across to a local machine if you're doing any large-scale harvesting. That'll make it easier for you to deal with corrupted files, broken downloads etc.

abc.net.au



<https://glam-workbench.github.io/web-archives/>

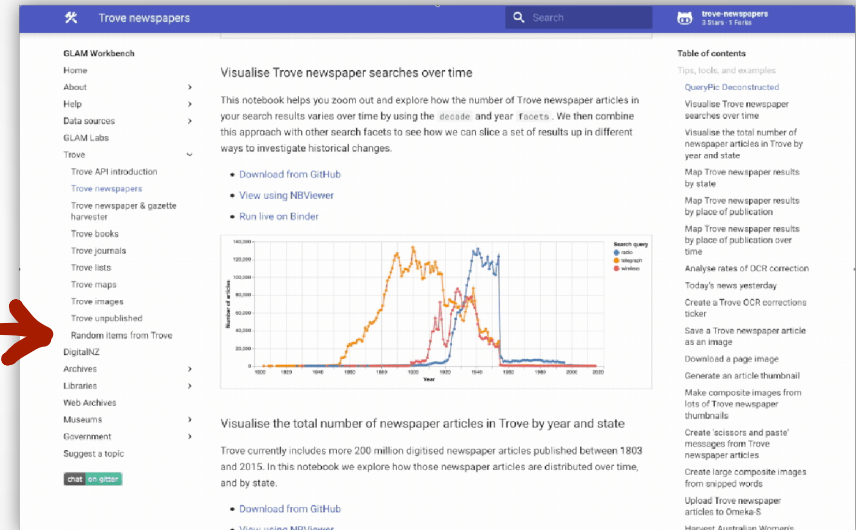
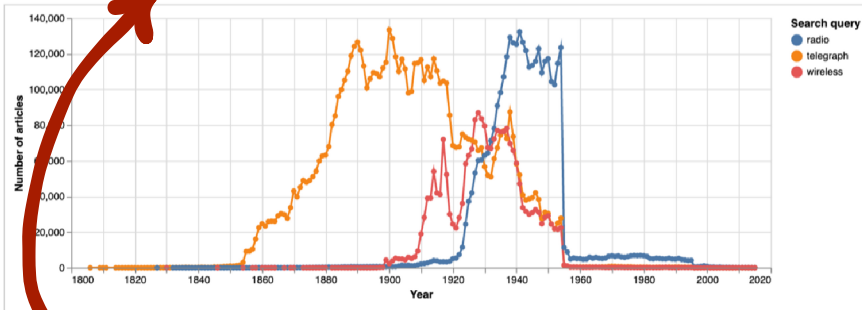
PLATFORMS  
&  
ENVIRONMENTS

# JUMP IN WITH BINDER

## Visualise Trove newspaper searches over time

This notebook helps you zoom out and explore how the number of Trove newspaper articles in your search results varies over time by using the `decade` and `year` facets. We then combine this approach with other search facets to see how we can slice a set of results up in different ways to investigate historical changes.

- Download from GitHub
- View using NBViewer
- Run live on Binder



<https://glam-workbench.net/using-binder/>

one click to start

## Using Reclaim Cloud

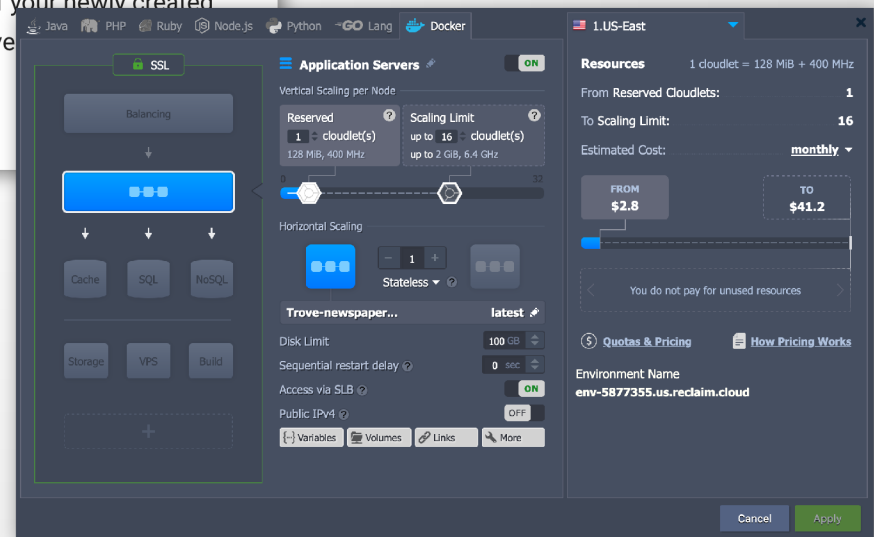
Launch on [Reclaim Cloud](#)

[Reclaim Cloud](#) is a paid hosting service, aimed particularly at supported digital scholarship in the humanities. Unlike Binder, the environments you create on Reclaim Cloud will save your data – even if you switch them off! To run this repository on Reclaim Cloud for the first time:

- Create a [Reclaim Cloud](#) account and log in.
- Click on the button above to start the installation process.
- A dialogue box will ask you to set a password, this is used to limit access to your Jupyter installation.
- Sit back and wait for the installation to complete!
- Once the installation is finished click on the 'Open in Browser' button of your newly created environment (note that you might need to wait a few minutes before even

See the [Using Reclaim Cloud](#) section GLAM Workbench [for more details.


# RECLAIM CLOUD



<https://glam-workbench.net/using-reclaim-cloud/>



## Using the Nectar Cloud



### GLAM Workbench

The GLAM Workbench is a collection of Jupyter notebooks to help researchers harvest, analyse, visualise, and share data from g...

[Details »](#)


[Add to Env](#) [Quick Deploy](#)

shared

The [Nectar Research Cloud](#) (part of the Australian Research Data Commons) provides cloud computing services to researchers in Australian and New Zealand universities. Any university-affiliated researcher can log on to Nectar and receive [up to 6 months of free cloud computing time](#). And if you need more, you can [apply for a specific project allocation](#).

The GLAM Workbench is available in the Nectar Cloud as a pre-configured application. This means you can get it up and going without worrying about the technical infrastructure – just fill in a few details and you're away! To create an instance of this repository,

# NECTAR CLOUD



### Configure Application: GLAM Workbench

Here you configure the GLAM Workbench itself by selecting a repository and creating a password. The password will be used to log in to your GLAM Workbench instance.

**Workbench \***

- ☒ DigitalNZ
- ☐ NAA RecordSearch
- ☐ Te Papa collections
- ☐ Trove newspapers
- ☐ Trove newspaper & gazette harvester
- ☐ Web Archives


**workbench:**

Select the GLAM Workbench repository you want to install. The options are from: <https://quay.io/organization/glamworkbench>.

**Password:**

Provide a password for this GLAM Workbench instance. It should be at least 7 characters, 1 capital letter, 1 digit and 1 special char, except single q...

[Show more](#)

**Confirm password** 

[Back](#) [Next](#)

<https://glam-workbench.net/using-nectar/>

A WORK IN PROGRESS

# HAVE SUGGESTIONS?




The screenshot displays the GitHub Issues page for the GLAM-Workbench repository. At the top, there are filters for '19 Open' and '3 Closed' issues, along with sorting options: Author, Label, Projects, Milestones, Assignee, and Sort. The issues are listed in descending order of age. Each issue entry includes a checkbox, a status icon (a green circle with a white dot), the issue title, labels (color-coded pills), the issue number, the time since it was opened, and the author's name. Some issues also show a comment icon with a number.

Issue #	Title	Labels	Opened	Author	Comments
#38	Use Docker images in Nectar cloud?	documentation, repository management	16 minutes ago	wragge	
#37	Documentation and testing for use on SWAN	documentation	19 minutes ago	wragge	
#36	Documentation for people/orgs wanting to add their own notebooks or repositories to the GLAM Workbench	documentation	29 minutes ago	wragge	1
#35	GLAM Name Index Search additions	enhancement	41 minutes ago	wragge	2
#34	Create help videos	documentation	1 hour ago	wragge	1
#33	Create a GLAM Workbench repository template	repository management	1 hour ago	wragge	
#32	Trove newspapers front page data	notebook idea	1 hour ago	wragge	
#31	RecordSearch web app?	notebook idea	1 hour ago	wragge	
#30	Update repos for automated Docker builds, Reclaim Cloud integration etc	documentation, repository management, to-do	1 hour ago	wragge	1
#29	Add citation file to GitHub repos	repository management, to-do	2 hours ago	wragge	1
#22	Add zenodo concept id to front matter in docs	documentation, to-do	on 8 Mar	wragge	3 tasks

<https://github.com/GLAM-Workbench/glam-workbench.github.io/issues>

# NEED HELP?

**OzGLAM Help**



**Welcome to OzGLAM Help!**


This is a place for people who love exploring the collections of Australia's GLAM organisations (that's Galleries, Libraries, Archives, and Museums). Ask a question if you're having problems, or simply share your latest discoveries. Here you'll find handy tips, news about recent developments, and links to useful tools. All GLAM users are welcome!


[Edit this banner >>](#)


GLAM Workbench ▾


Latest

Top

 Edit

 New Topic




 Topic


Replies

Views

Activity

 About the GLAM Workbench category

Get help with the GLAM Workbench's growing collection of digital tools, tutorials, and pre-harvested datasets.




0

51

Sep '20

New notebook – finding non-English language newspapers in Trove

trove




0

84

18 Jan

New dataset and notebooks – twenty years of ABC Radio National

trove




0

97

18 Jan

Using web archives

trove nationallibrarynz ukwebarchive





0

44

Nov '20

Trove Newspaper Harvester updated – harvest text from the Australian Women's Weekly!




2

164

Nov '20

<https://ozglam.chat/c/glam-workbench/>

# NEED HELP?



**GLAM Workbench:**  
Explore data from  
libraries, archives &  
museums


2 videos • 9 views • Updated 2 days ago


Public ▾


✂ ➡ ...

The GLAM Workbench is collection of tools, tutorials, examples, and hacks to help you work with data from galleries, libraries, archives, and museums (the GLAM sector). The videos here will help you make use of the GLAM Workbench.

See: <https://glam-workbench.net/>

 SORT

 **GLAM Workbench – use QueryPic to visualise searches in Trove's digitised newspapers**  
Tim Sherratt  
12:50

 **GLAM Workbench – use QueryPic to visualise searches in Trove's digitised newspapers (part 2)**  
Tim Sherratt  
16:09

[https://www.youtube.com/playlist?list=PLAclcciEeCD2z2BWQ2r3xD\\_Q8c05HppfP](https://www.youtube.com/playlist?list=PLAclcciEeCD2z2BWQ2r3xD_Q8c05HppfP)



# NEED HELP?

GLAM Workbench

Home

About

Help

Data sources

Trove

DigitalNZ

Archives

Libraries

Web Archives

Museums

Government

Suggest a topic

Ask a question

Book a chat

Latest news!

♥ Sponsor

Welcome to the GLAM Workbench

Winner of the British Library Labs Research Award 2020!

Here you'll find a collection of tools, tutorials, examples, and hacks to help you work with data from galleries, libraries, archives, and museums (the GLAM sector). The primary focus is Australia and New Zealand, but new collections are being added all the time. Let me know if the GLAM Workbench might help your research? Or perhaps you've struck a problem. Book in for a chat and I'll see what I can do to help!

Table of contents

Quick start

Finding GLAM

Harvesting


Data sources

Asking different

Hacking her

Bringing do

Do I need to



Tim Sherratt

GLAM Workbench chat

🕒 30 min

💬 Web conferencing details provided upon confirmation.

Are you interested in discussing how the [GLAM Workbench](#) might help your research? Or perhaps you've struck a problem. Book in for a chat and I'll see what I can do to help!

Select a Date & Time

July 2021

MON	TUE	WED	THU	FRI	SAT	SUN
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

🌐

 Sydney, Melbourne Time (3:32pm) ▾

Powered by Calendly

# THE GLAM WORKBENCH IS...

- Free!
- Openly licensed to encourage reuse & modification
- Shared through GitHub
- Preserved in Zenodo

HELP TRAPPED. Inside trove, cannot Escape

Optical CHARACTER Recognition HAS achieved consciousness

# THANK YOU!



**Jim Groom**  
@jimgroom



💖 Reclaim Hosting is sponsoring @wragge's GLAM Workbench because it's a remarkable project that has the telltale signs of a selfless labor of love.



Sponsor @wragge on GitHub Sponsors  
Historian and hacker who builds things that help people use, analyse, and explore the digital collections of ...  
[github.com](https://github.com)

9:33 PM · Sep 28, 2021 · Twitter Web App

19 sponsors are funding wragge's work.



**@wragge's goal is to have 50 sponsors**

I make things for people to use. Knowing that 50 people valued my work enough to sponsor it would give me a huge boost and help me argue for the importance of this sort of work.



<https://github.com/sponsors/wragge>

TIM SHERRATT

@wragge

<https://glam-workbench.net/>