Website collecting by the State Library of NSW including Pandora

Brendan Somes and Boris Nazarov

12 March 2024

STATE LIBRARY®
NEW SOUTH WALES

Legislation – Library Act 1939

- Establishment legislation for the Library
- Updated in 2019 to enable deposit of digital publications as well as print publications
- Previously Copyright Act 1879 only supported print legal deposit

STATE LIBRARY®
NEW SOUTH WALES

Legislation – Library Act 1939

- https://legislation.nsw.gov.au/view/html/inforce/current/act-1939-040#pt.3A
- To enable website collecting have section

Collection of freely available relevant library material

STATE LIBRARY®
NEW SOUTH WALES

Legislation – NSW Government Publications

- C2022-02 Deposit of New South Wales Government Publications requires deposit of print and digital publications.
- Released March 2022 replacing 2000 Premier's Memorandum

STATE LIBRARY®
NEW SOUTH WALES

Legislation – NSW Government Publications

- Treasury issued in August 2023
- TPG23-10 Annual Reporting Requirements
- This addresses the deposit of NSW Government Annual Reports
- Replaced M2013-09 Production Costs of Annual Reports

STATE LIBRARY®
NEW SOUTH WALES

Legislation – Australian Government Publications

- Commonwealth Library Deposit and Free Issue Schemes
- Requirements for deposit of Australian Government print and digital publications

STATE LIBRARY®
NEW SOUTH WALES

Collecting Policies

- Collection Development Policy
- https://www.sl.nsw.gov.au/sites/default/files/collection_development_policy.pdf
- *Materials published in New South Wales are collected under the legal and government deposit provisions of the Library Amendment Act 2019 (NSW), the Circular C2022-02 Deposit of New South Wales Government Publications and the Premier's Memorandum M2013-09 Production Costs of Annual Reports*.

STATE LIBRARY®
NEW SOUTH WALES

Collecting Digital Publications - NED

- NED is National edeposit service
- Established 2019 by the National, State and Territory Libraries of Australia
- Enables publisher to deposit electronic books, serials, maps and sheet music
- www.ned.gov.au

STATE LIBRARY®
NEW SOUTH WALES

Collecting Digital Publications – Archive-It

- Archive-It is an Internet Archive tool the Library uses to collect lots of websites all at once – for example – nsw.gov.au domain harvests
- https://collection.sl.nsw.gov.au/record/74VKOeJgbZp3
- https://www.archive-it.org/collections/4661

STATE LIBRARY®
NEW SOUTH WALES

Collecting Digital Publications – Pandora

- Pandora is the website collecting tool developed by the National Library of Australia
- Use to collect individual websites
- Partners - https://pandora.nla.gov.au/partners.html

STATE LIBRARY®
NEW SOUTH WALES

Collecting Digital Publications – Pandora

- Australian Web Archive - https://www.nla.gov.au/collections/building-our-collections/australian-web-archive

STATE LIBRARY®
NEW SOUTH WALES

Collecting Digital Publications – Vizie

- Vizie is a CSIRO social media collecting tool.
- Library has been using for a number of years
- Collects social media posts from a number of platforms – Facebook, Instagram, Twitter, Reddit
- https://socialmediaarchive.sl.nsw.gov.au/prototype/live.html

STATE LIBRARY®
NEW SOUTH WALES

Collecting Digital Publications – Work Ahead

- Whole of Domain v Single Website
- .au harvest and other harvests
- Social Media
- Government Publications
- Technical challenges …

STATE LIBRARY®
NEW SOUTH WALES

Introduction

- PANDORA (Preserving and Accessing Networked Documentary Resources of Australia)
- Archving content, not replicating functionality.

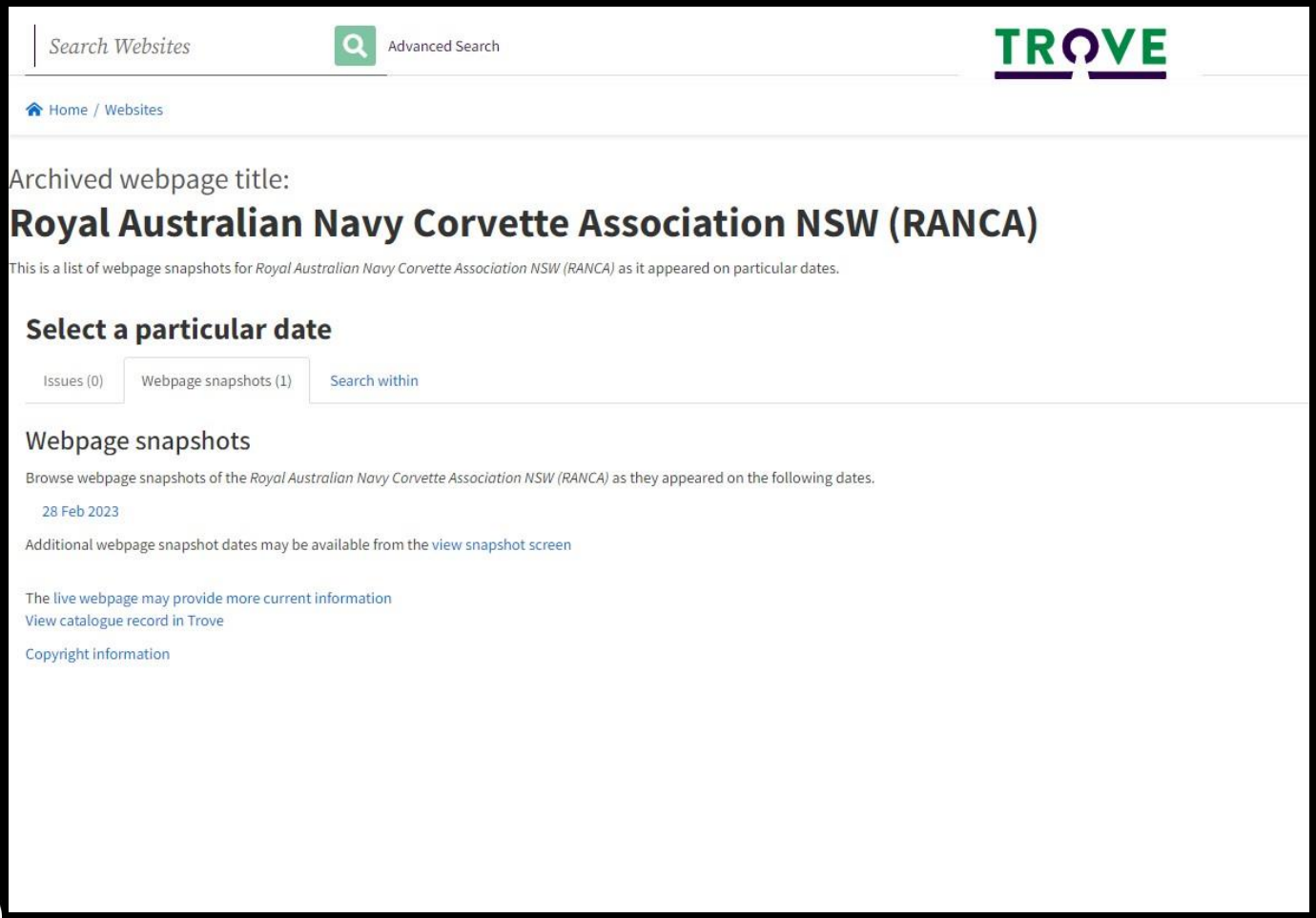
STATE LIBRARY®
NEW SOUTH WALES

What can PANDORA collect?

- Self-contained sites with text, images, documents, with few or no links to external sites.
- Sites that change over time
  eg annual festivals
- Collections of multiple sites around a theme/event eg elections

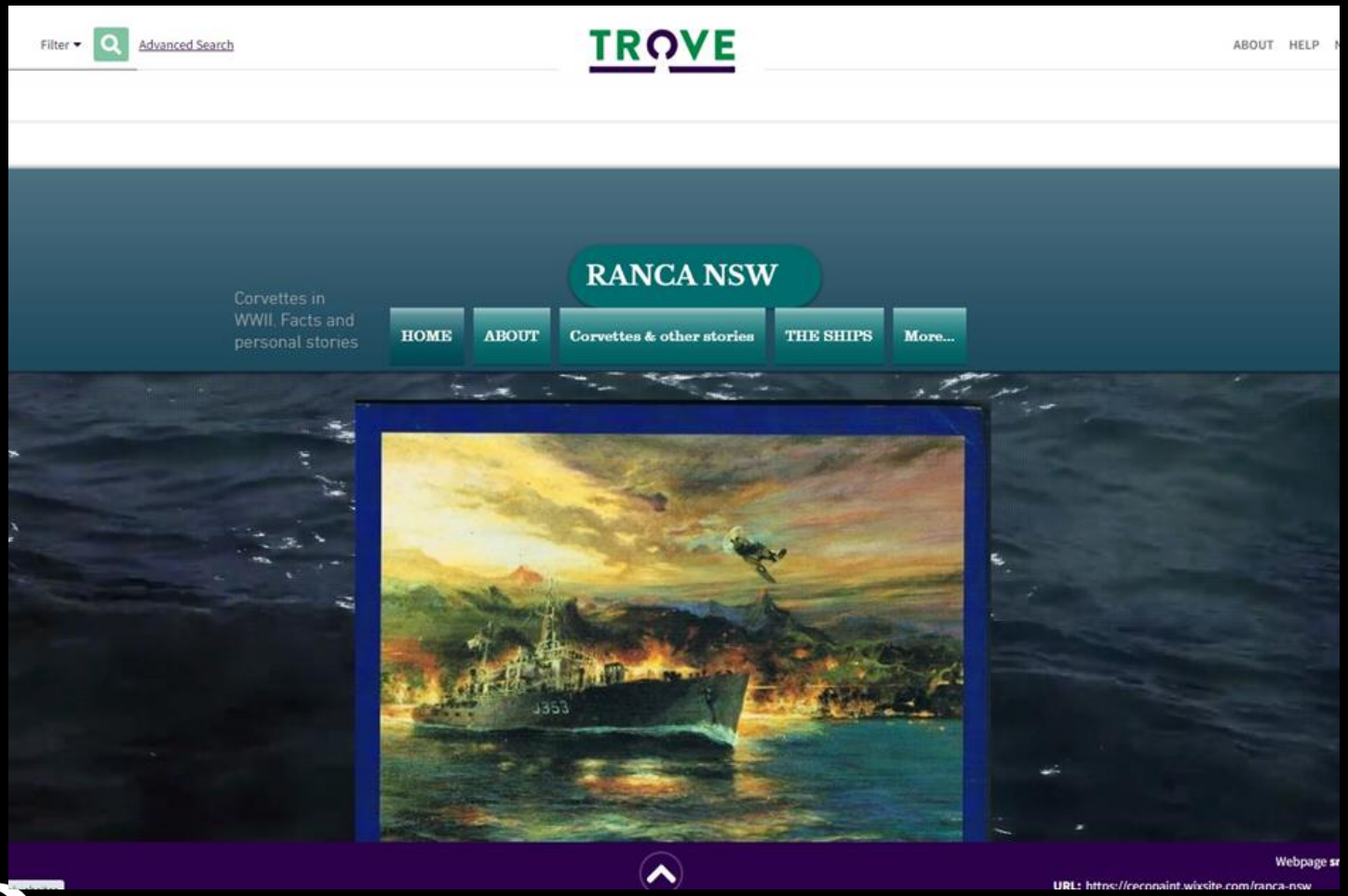
STATE LIBRARY®
NEW SOUTH WALES

Self-contained sites with text, images, documents



STATE LIBRARY®
NEW SOUTH WALES

RANCA



STATE LIBRARY®
NEW SOUTH WALES

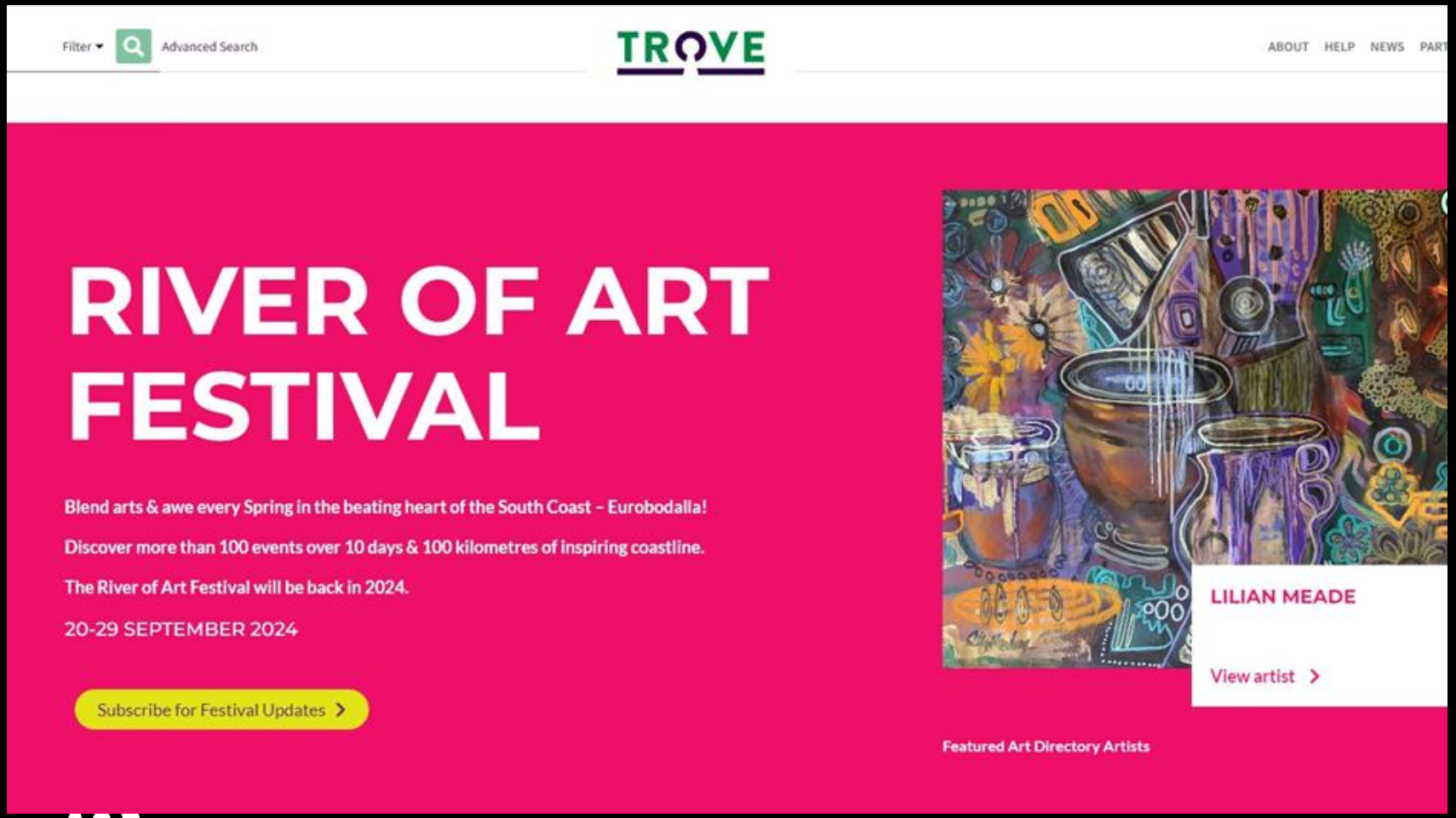
Single sites with text, images, documents

STATE LIBRARY®
NEW SOUTH WALES

Synthetics

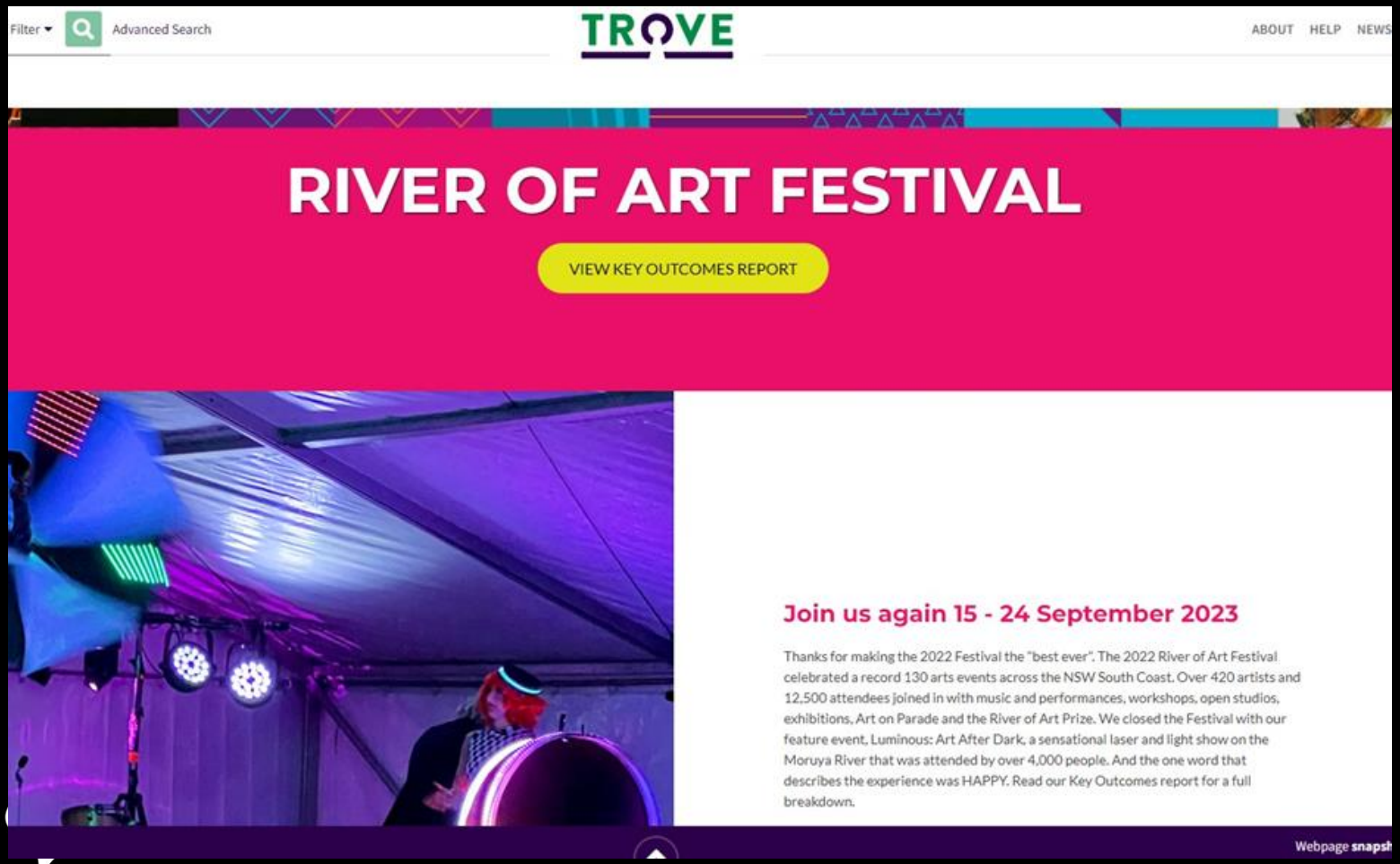
STATE LIBRARY®
NEW SOUTH WALES

Sites that change over time (2024)



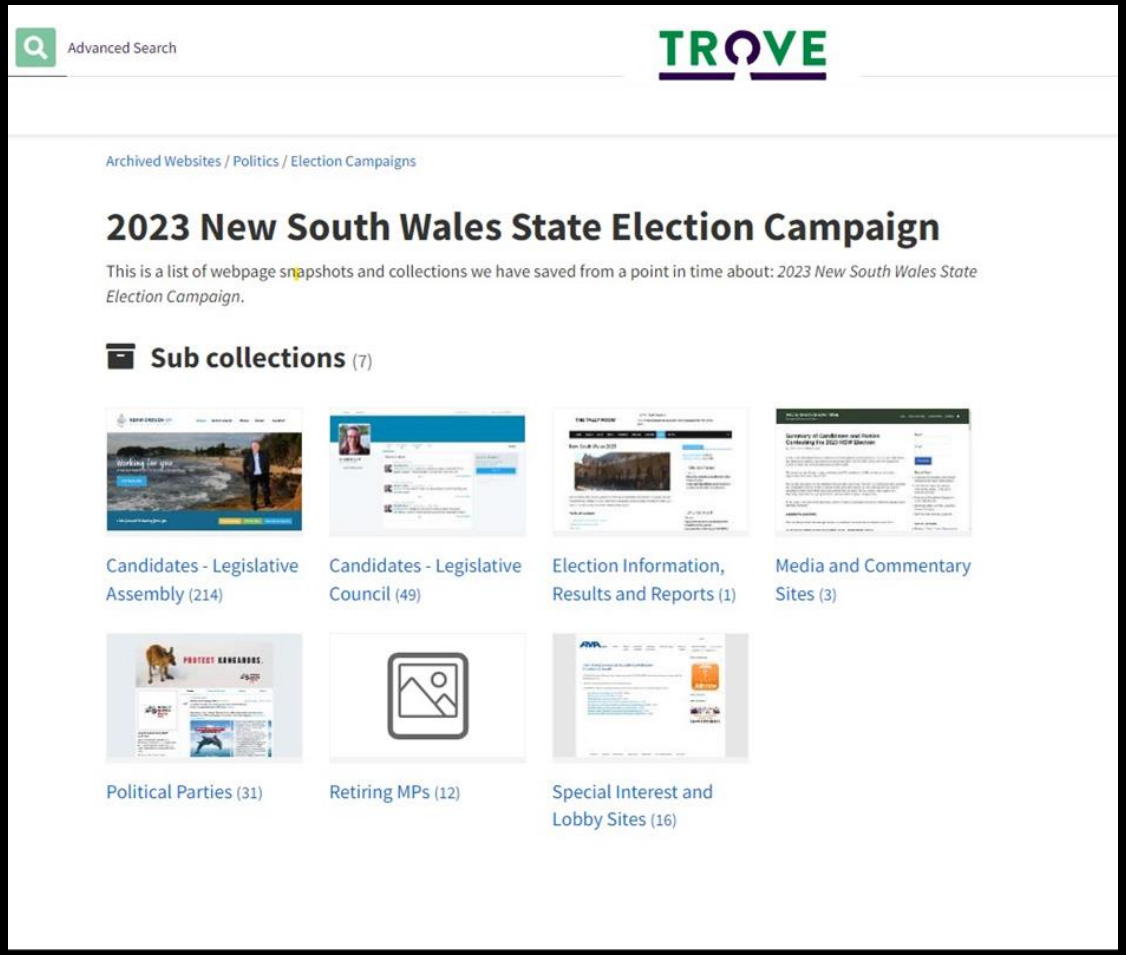
STATE LIBRARY®
NEW SOUTH WALES

Eurobodalla River of Art Festival (2023)



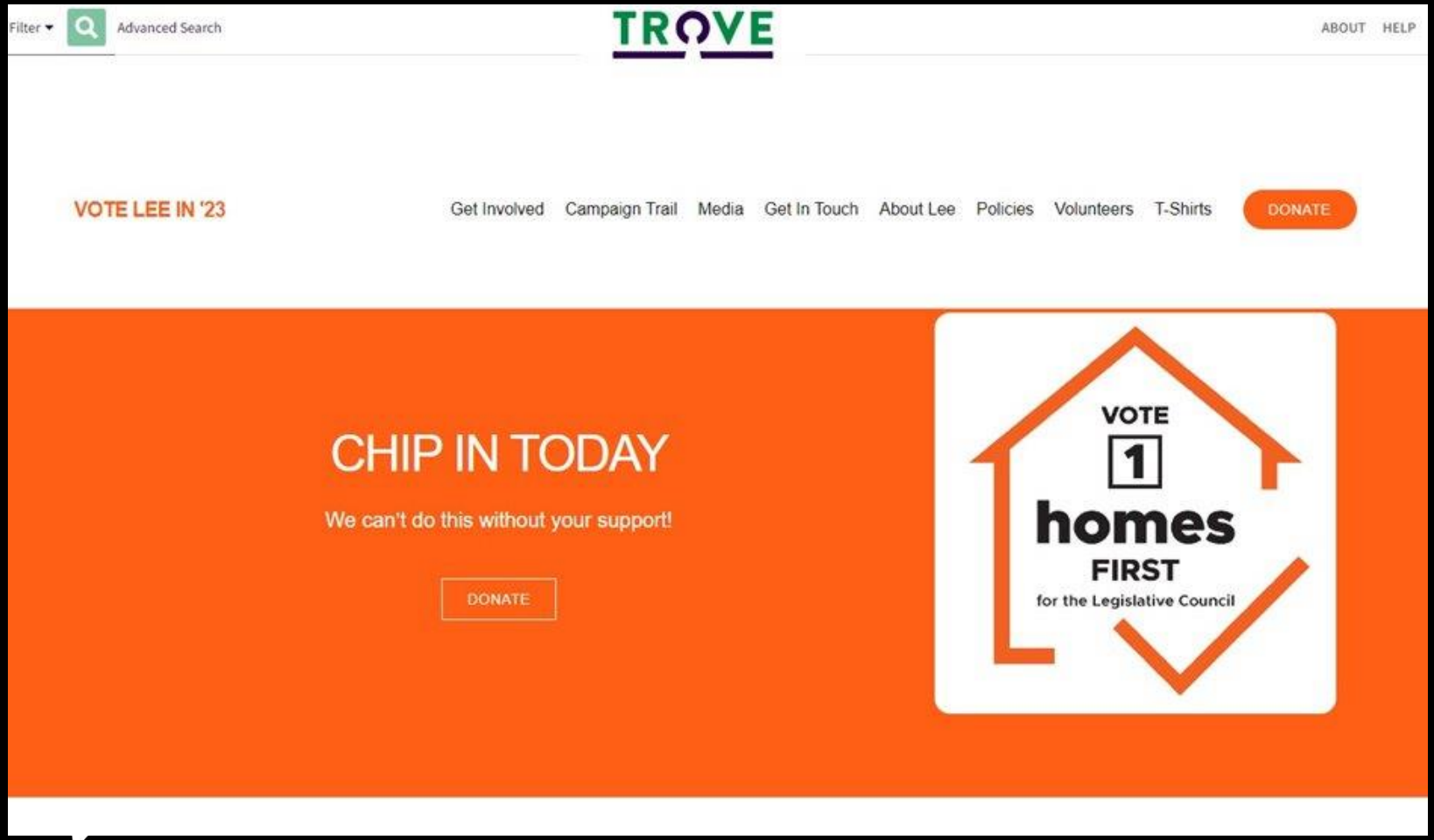
STATE LIBRARY®
NEW SOUTH WALES

Collections of multiple sites

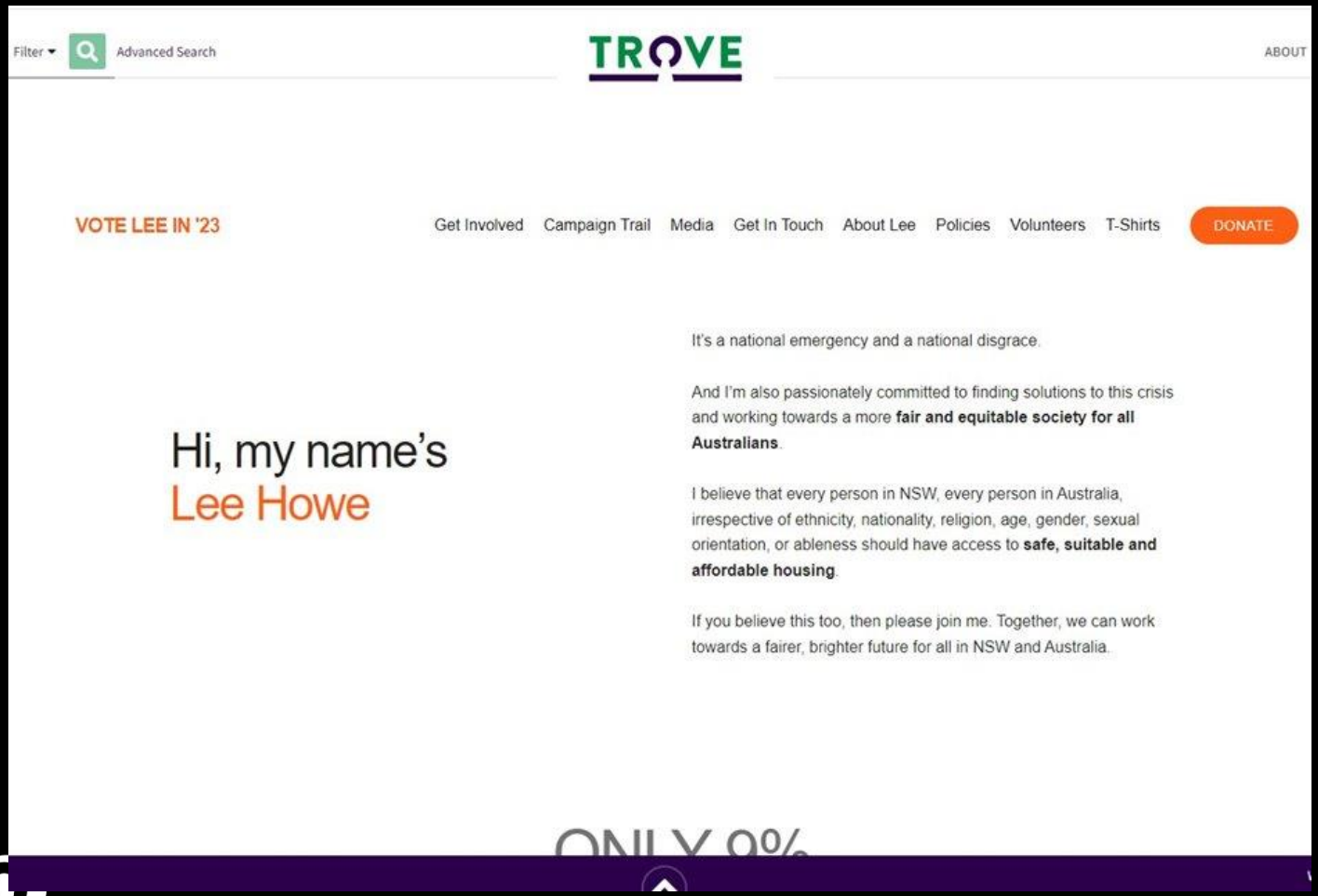


STATE LIBRARY®
NEW SOUTH WALES

Candidate website / Lee Howe

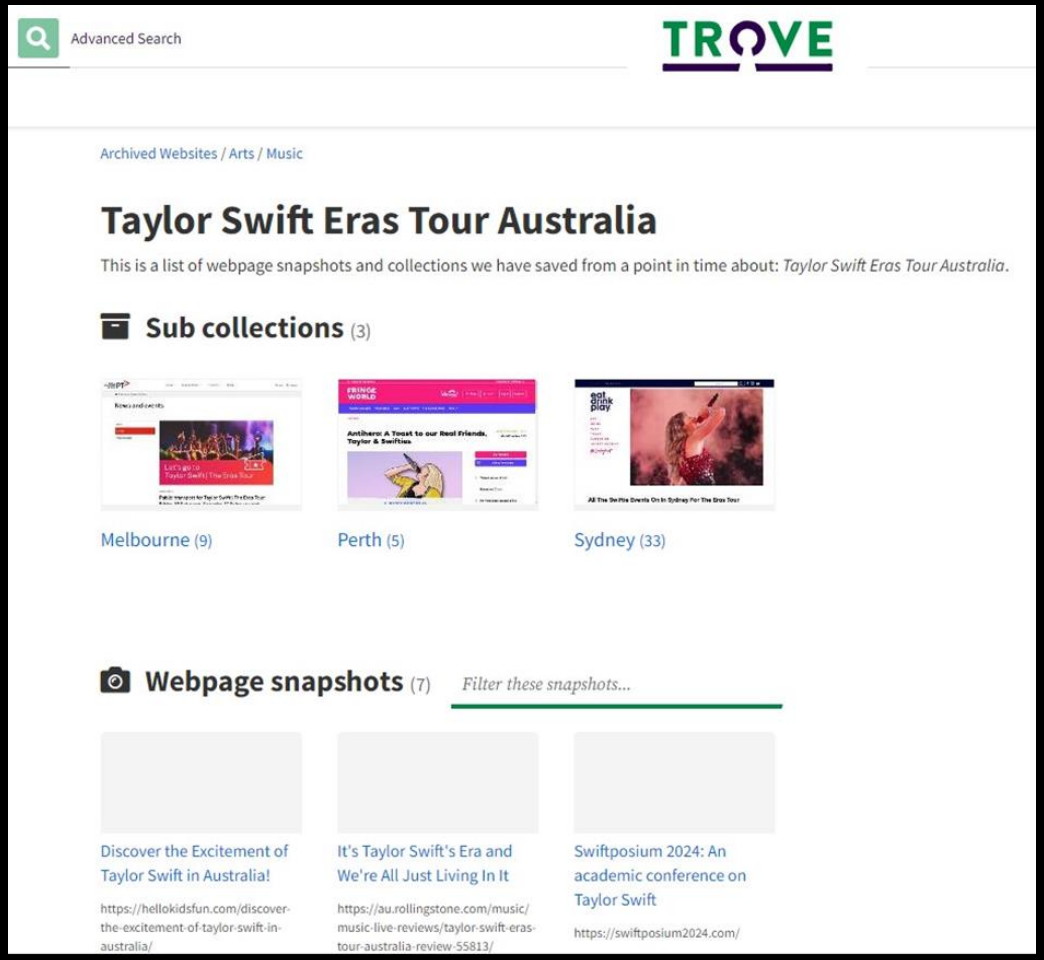


Candidate website / Lee Howe



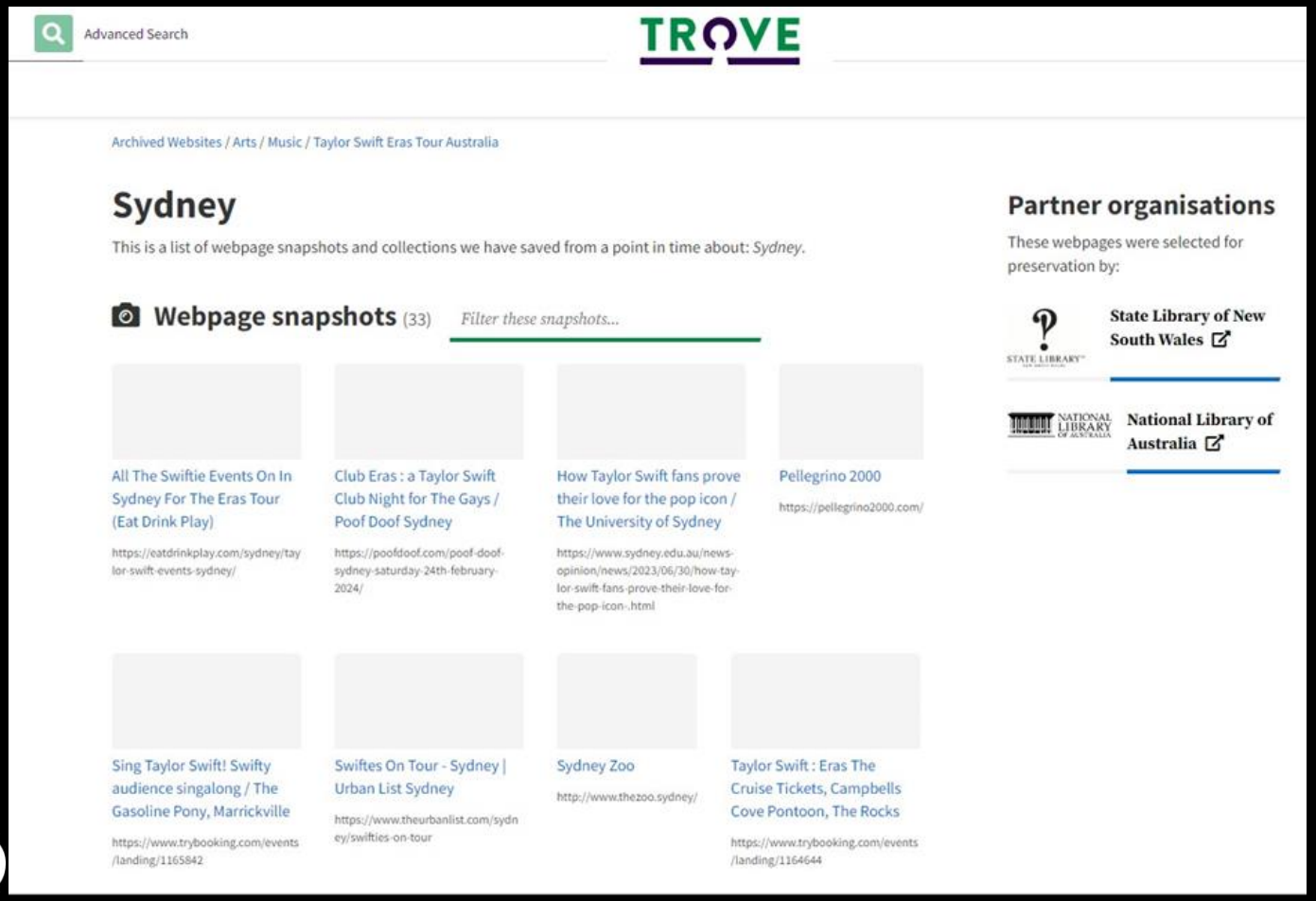
STATE LIBRARY®
NEW SOUTH WALES

Taylor Swift



STATE LIBRARY®
NEW SOUTH WALES

Taylor Swift / Sydney



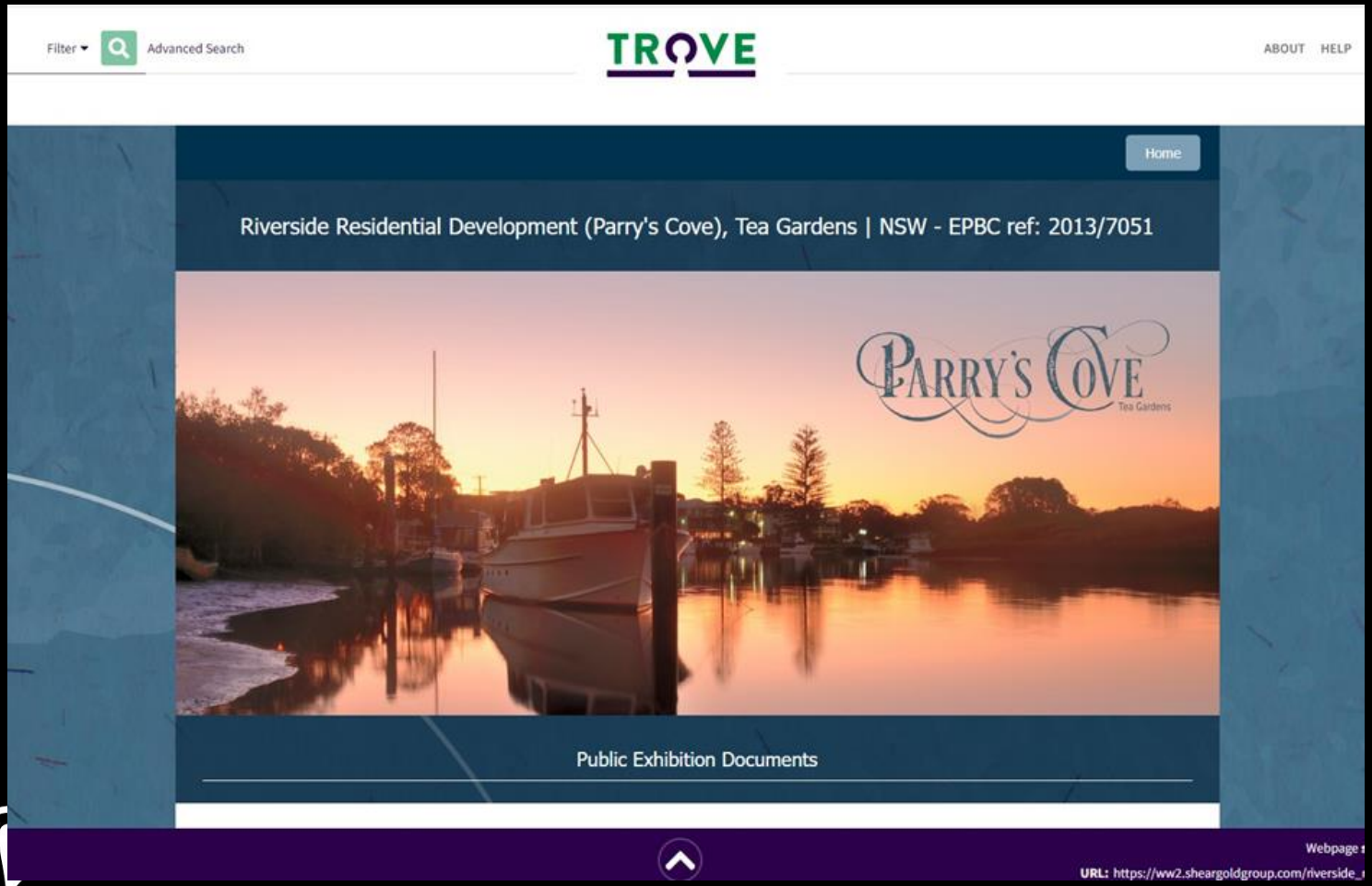
STATE LIBRARY®
NEW SOUTH WALES

What PANDORA has issues collecting or not suitable for collecting

- Single PDFs (-> NED)
- Websites that use search functionality to locate webpages
- Social media, no Facebook, Instagram, Twitter (now X)
- Some videos, hosted videos possibly, but not linked videos eg Youtube
- Some commercial news sites eg news.com.au will block harvesting
- Paywalled or sites requiring logins.

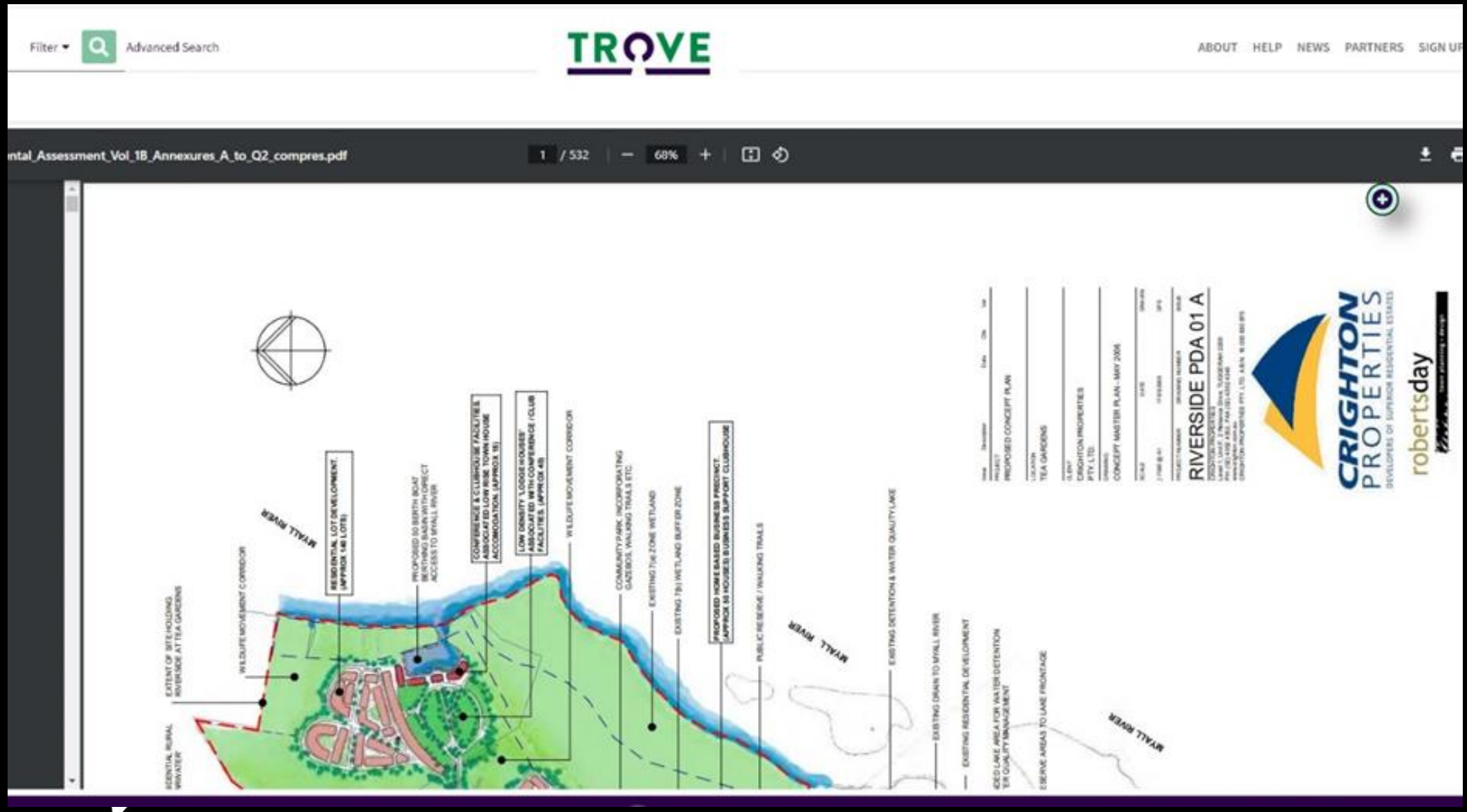
STATE LIBRARY®
NEW SOUTH WALES

Parry's Cove / Residential Development (site with multiple PDFs)



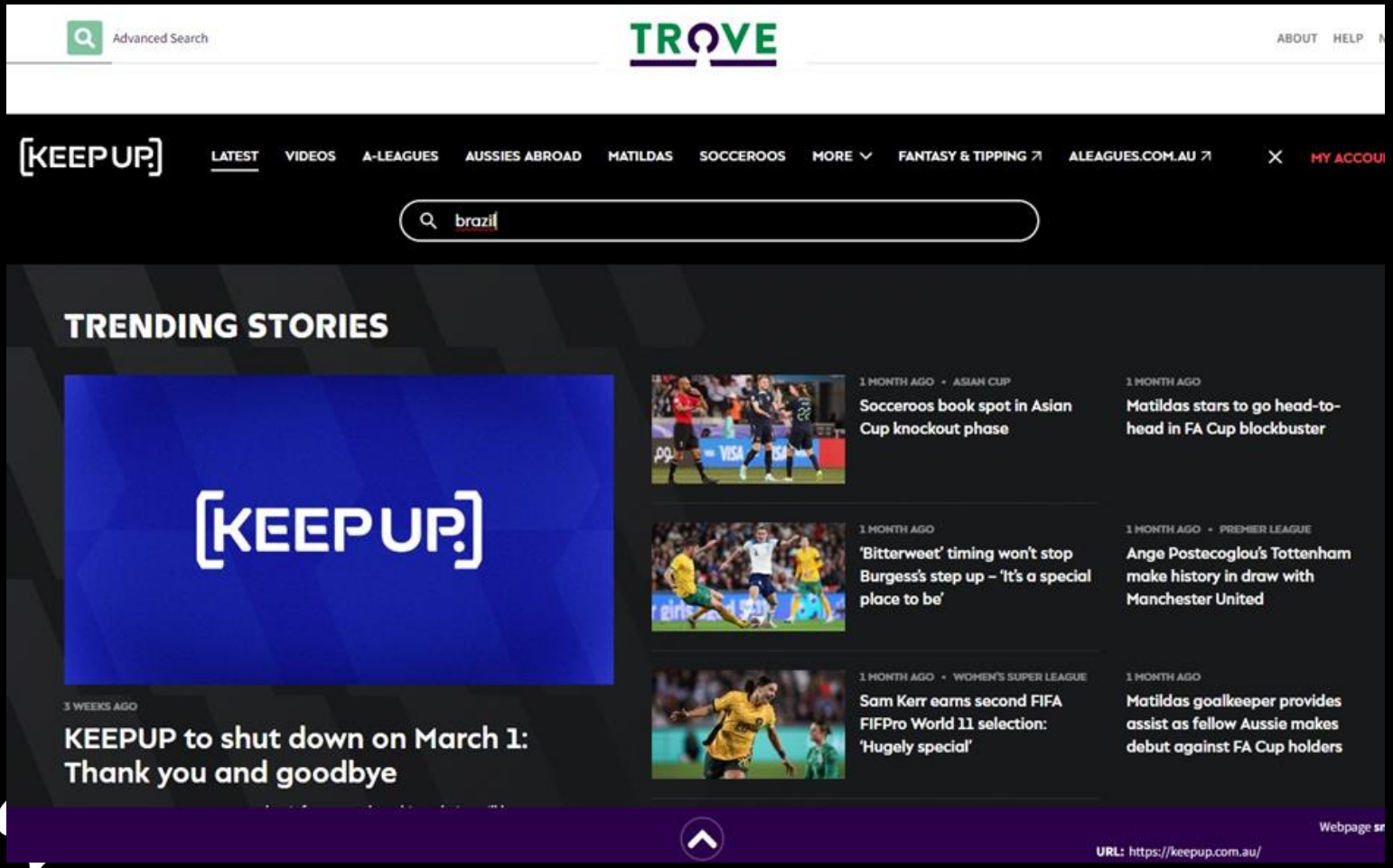
STATE LIBRARY®
NEW SOUTH WALES

Single PDFs (-> NED)



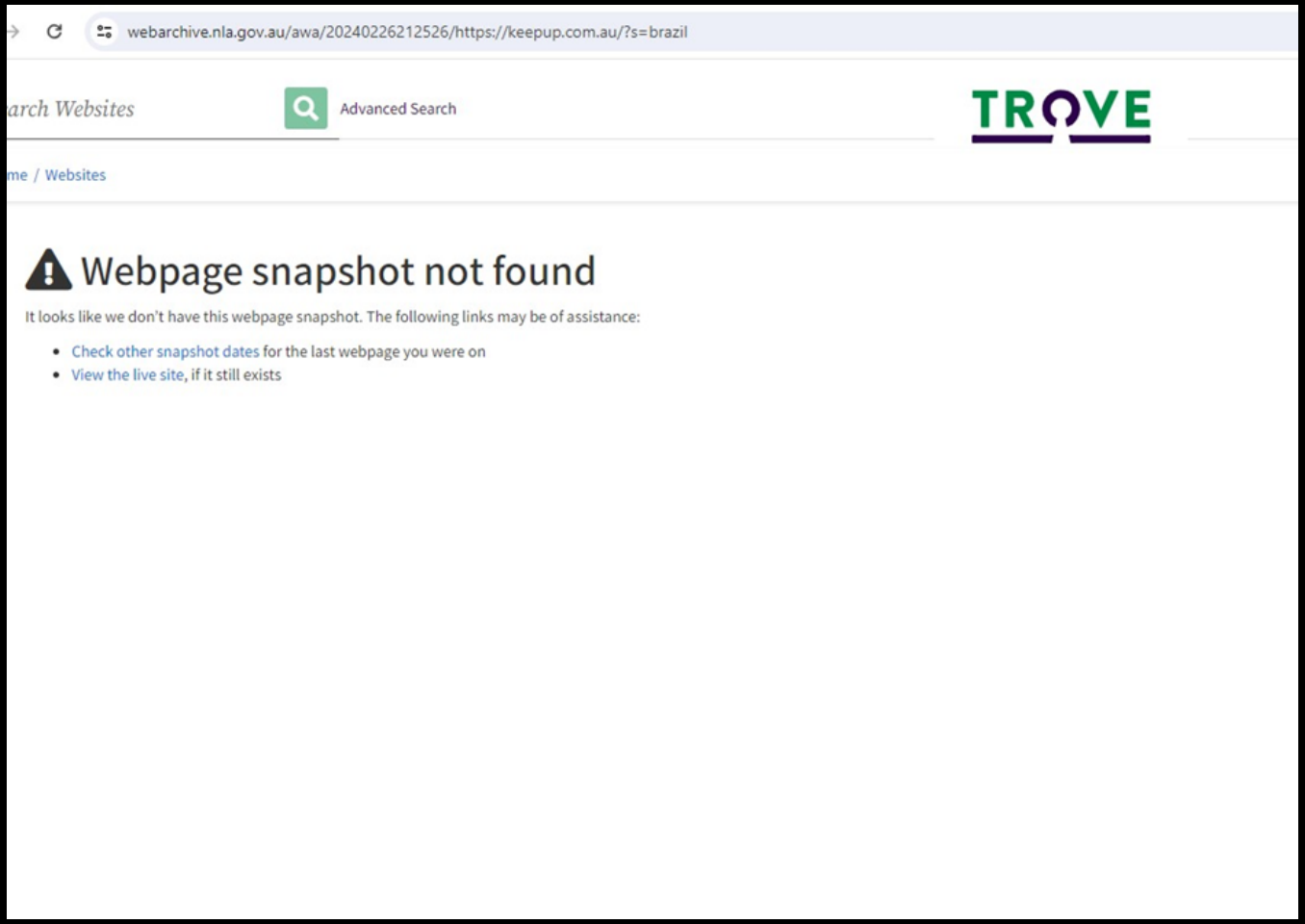
STATE LIBRARY®
NEW SOUTH WALES

Websites with search boxes



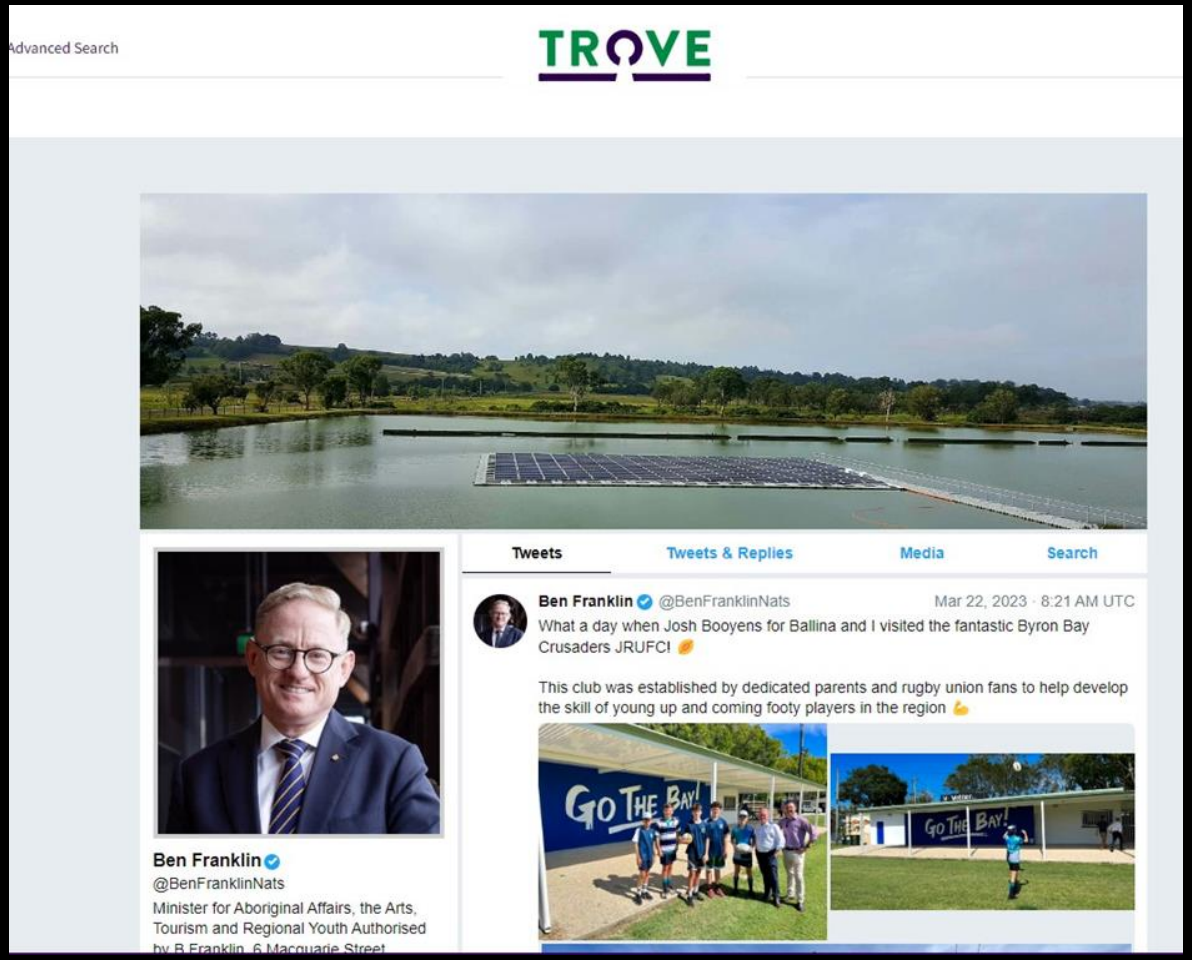
STATE LIBRARY®
NEW SOUTH WALES

KeepUp



STATE LIBRARY®
NEW SOUTH WALES

Twitter (via Nitter), no longer working since July 2023.



STATE LIBRARY®
NEW SOUTH WALES

PANDORA – Basic Workflow

- Selection
- PANDAS (PANDORA Digital Archiving System)
- Scheduling (daily, weekly, monthly etc...) options
- Quality checking eg images and links etc...
- Available in Trove
- Libraries Australia (links back to PANDAS)
- SLNSW catalogue

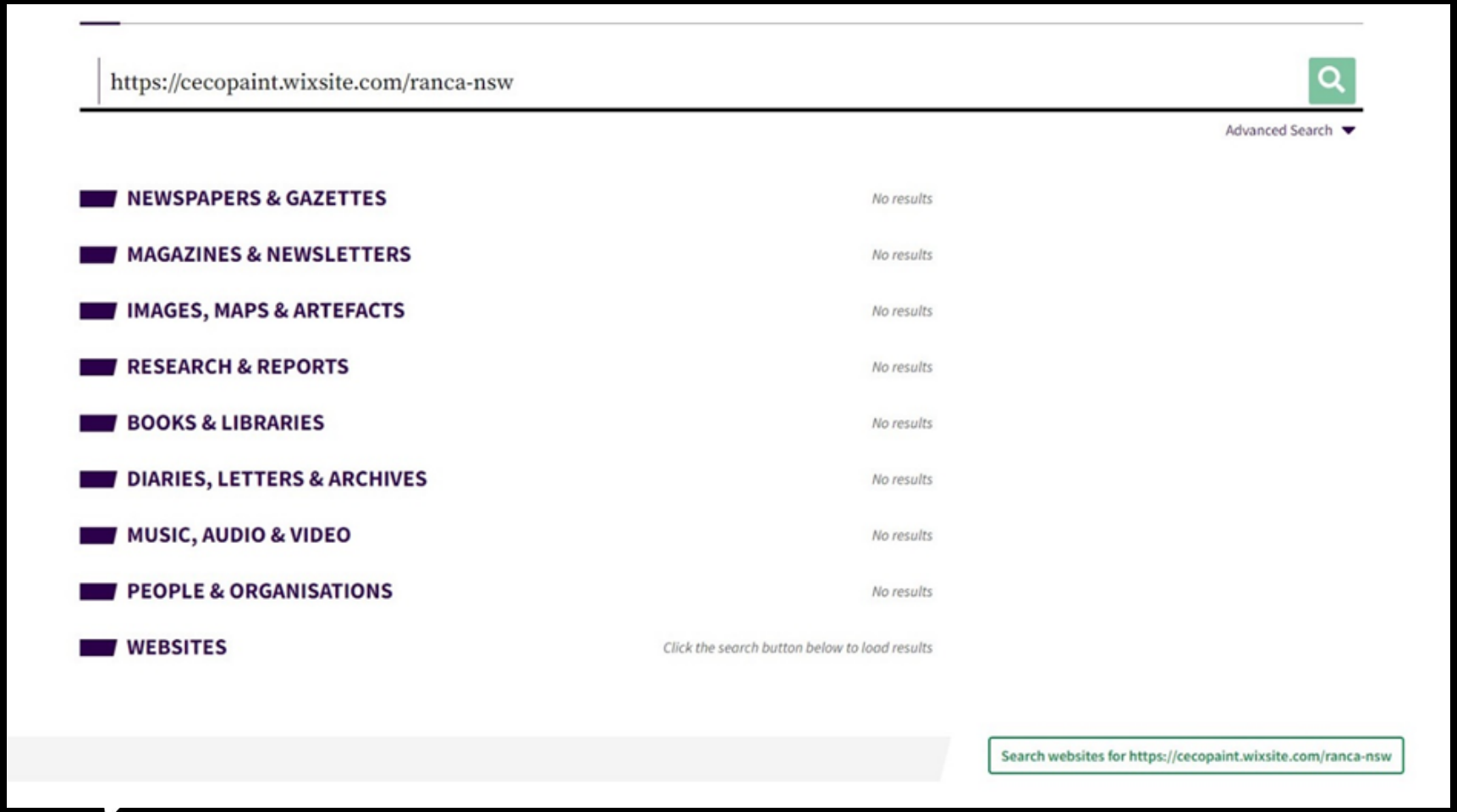
STATE LIBRARY®
NEW SOUTH WALES

PANDORA – How to find in Trove

- Search with URL
- Search with keywords

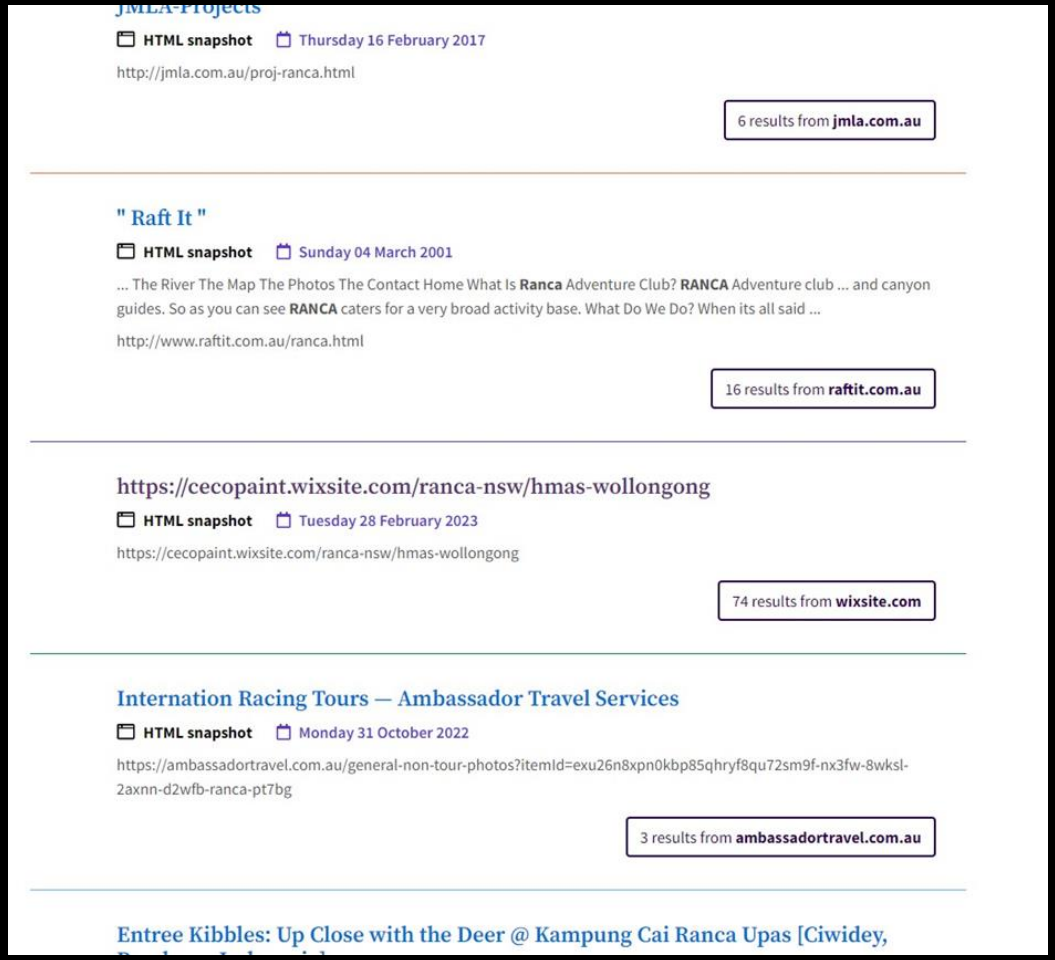
STATE LIBRARY®
NEW SOUTH WALES

With URL



STATE LIBRARY®
NEW SOUTH WALES

With keyword



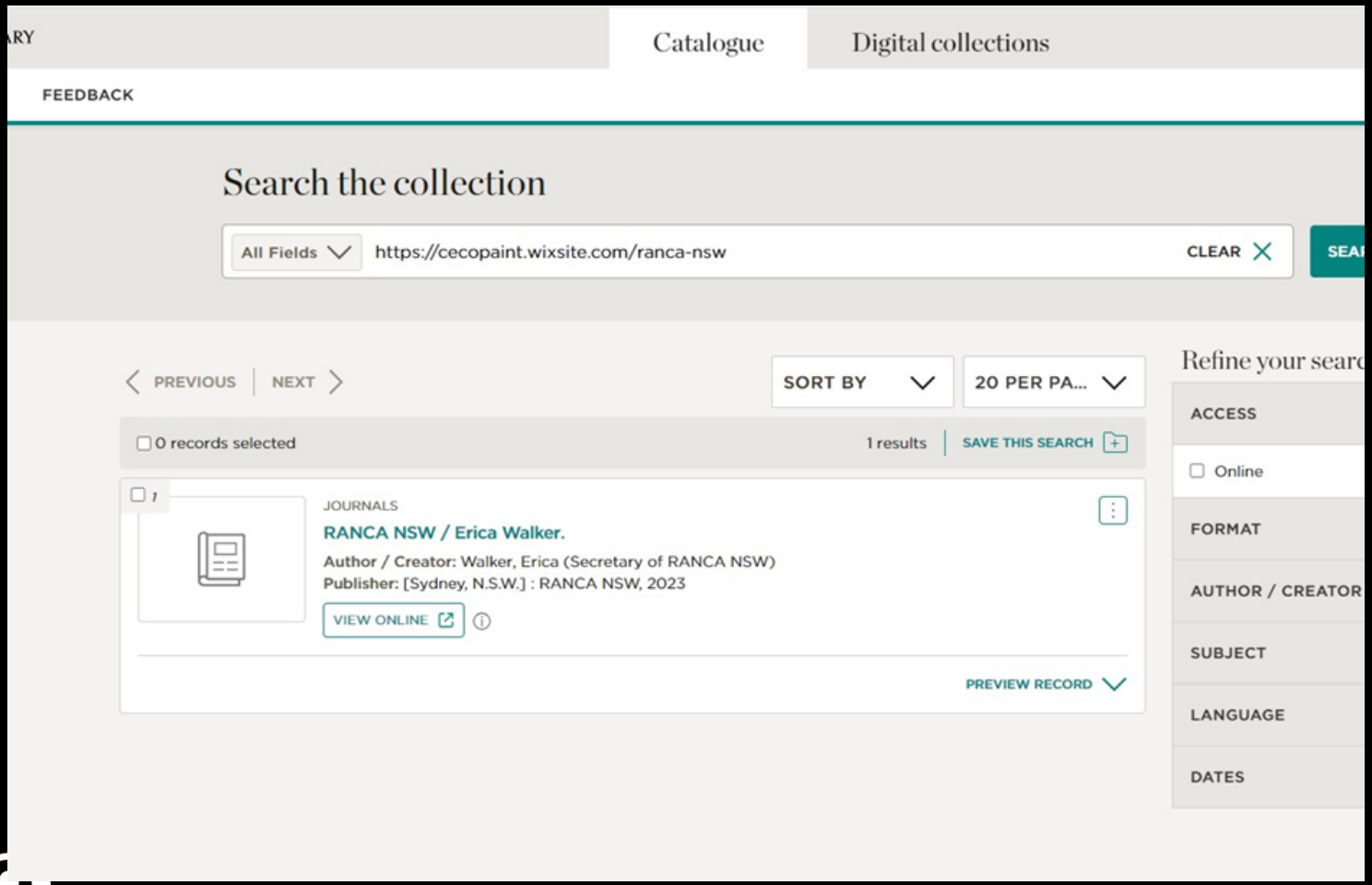
STATE LIBRARY®
NEW SOUTH WALES

PANDORA – How to find in SLNSW Catalogue

- Search with URL
- Search with keywords
- Link to Trove in SLNSW catalogue record

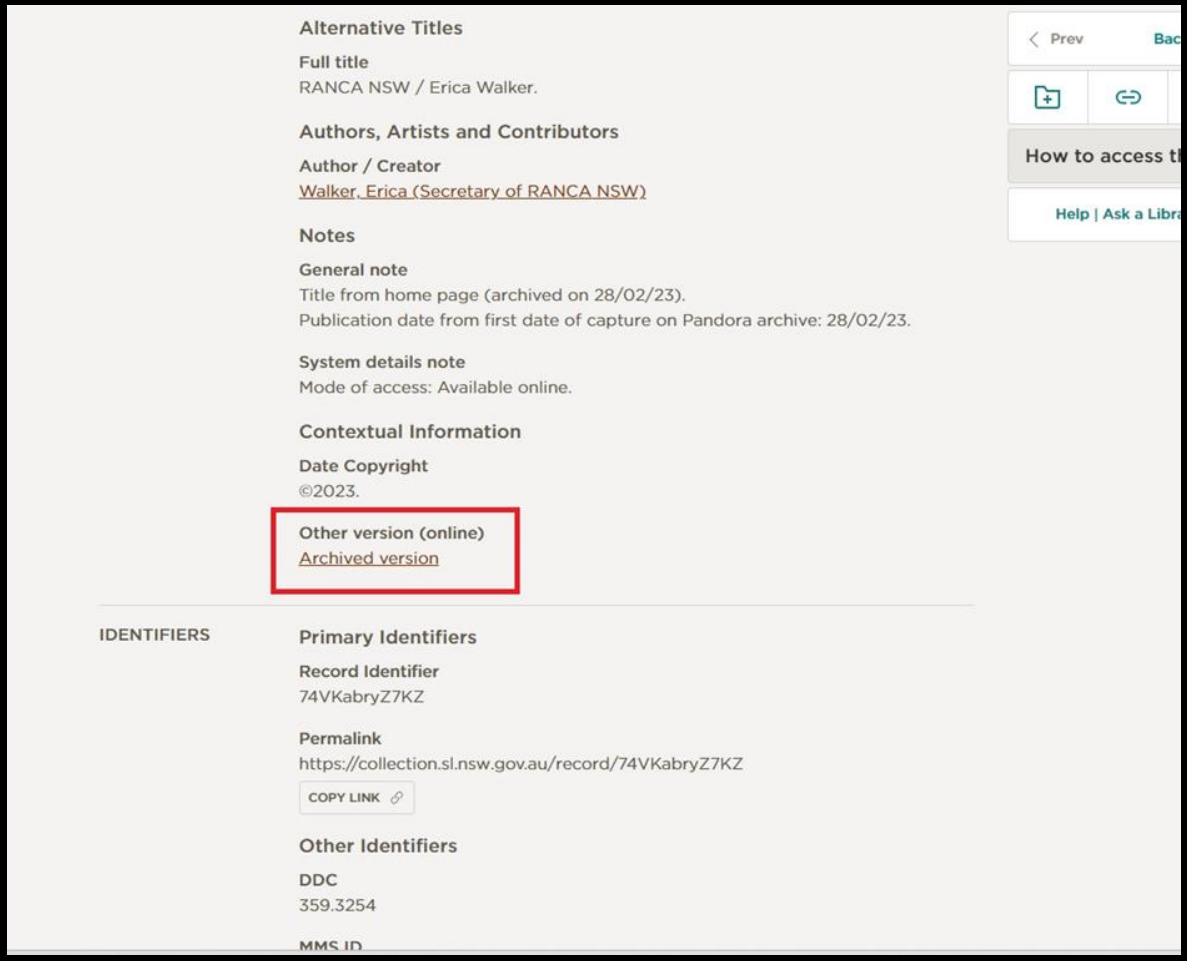
STATE LIBRARY®
NEW SOUTH WALES

SLNSW search



STATE LIBRARY®
NEW SOUTH WALES

SLNSW catalogue record



Alternative Titles

Full title
RANCA NSW / Erica Walker.

Authors, Artists and Contributors

Author / Creator
Walker, Erica (Secretary of RANCA NSW)

Notes

General note
Title from home page (archived on 28/02/23).
Publication date from first date of capture on Pandora archive: 28/02/23.

System details note
Mode of access: Available online.

Contextual Information

Date Copyright
©2023.

Other version (online)
Archived version

IDENTIFIERS

Primary Identifiers

Record Identifier
74VKabryZ7KZ

Permalink
https://collection.sl.nsw.gov.au/record/74VKabryZ7KZ

COPY LINK

Other Identifiers

DDC
359.3254

MMS ID

Prev

How to access t

Help | Ask a Libra

STATE LIBRARY®
NEW SOUTH WALES

Final

- Although not all websites can be archived due to issues mentioned in previous slides, it is worth attempting as there are different gathering methods and options.
- Send any suggestions to

collections@sl.nsw.gov.au

STATE LIBRARY®
NEW SOUTH WALES

Mitchell
printed books
card catalogue
Also check
the computer
catalogues