Website collecting by the State Library of NSW including Pandora

Brendan Somes and Boris Nazarov

12 March 2024



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



Legislation – Library Act 1939

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

Collection of freely available relevant library material



Legislation – NSW Government Publications

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



Legislation – NSW Government Publications

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



Legislation – Australian Government Publications

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



Collecting Policies

- Collection Development Policy
- <u>https://www.sl.nsw.gov.au/sites/default/file</u> <u>s/collection_development_policy.pdf</u>
- 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.



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



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
- <u>https://collection.sl.nsw.gov.au/record/74V</u>
 <u>KOeJgbZp3</u>
- <u>https://www.archive-it.org/collections/4661</u>



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



Collecting Digital Publications – Pandora

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



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
- <u>https://socialmediaarchive.sl.nsw.gov.au/pr</u>
 <u>ototype/live.html</u>

Collecting Digital Publications – Work Ahead

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



Introduction

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



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

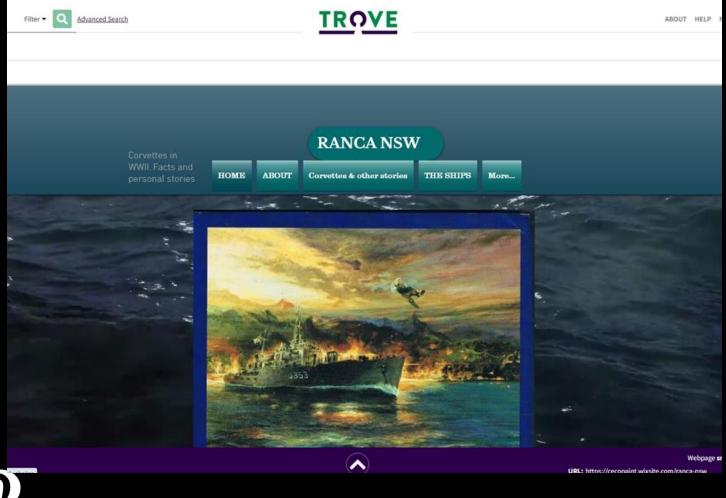


Self-contained sites with text, images, documents

Home / Websites rchived webpage title: Royal Australian Navy Corvette Association NS Select a particular date	
oyal Australian Navy Corvette Association NSW (RANCA) as it appeared on particular date	
s is a list of webpage snapshots for Royal Australian Navy Corvette Association NSW (RANCA) as it appeared on particular date	
is a list of webpage snapshots for Royal Australian Navy Corvette Association NSW (RANCA) as it appeared on particular date	
select a particular date	
Issues (0) Webpage snapshots (1) Search within	
Vebpage snapshots	
rowse webpage snapshots of the Royal Australian Navy Corvette Association NSW (RANCA) as they appeared on the following	g dates.
28 Feb 2023	
dditional webpage snapshot dates may be available from the view snapshot screen	
he live webpage may provide more current information iew catalogue record in Trove	
opyright information	

STATE LIBRARY®

RANCA



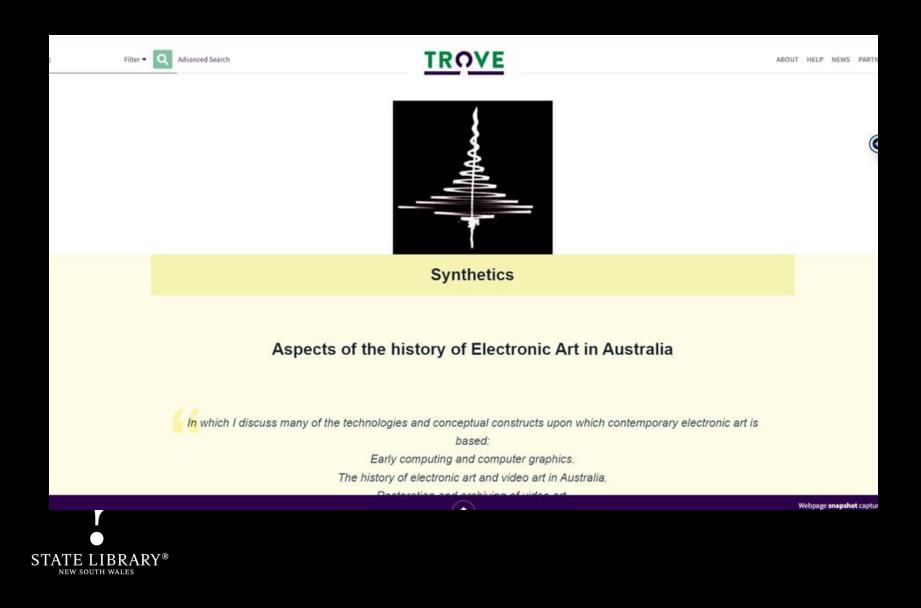


Single sites with text, images, documents

Search Websites	Advanced Search	TROVE
A Home / Websites		
Archived webpage title:		
Synthetics: Aspe	ects of the history of Elec	tronic Art in Australia
This is a list of webpage snapshots for Synthematic	tics: Aspects of the history of Electronic Art in Australia as it appeared	on particular dates.
Select a particular da	te	
Issues (0) Webpage snapshots (1)	Search within	
Webpage snapshots		
Browse webpage snapshots of the Syntheti	cs: Aspects of the history of Electronic Art in Australia as they appeared	ed on the following dates.
05 Mar 2024		
Additional webpage snapshot dates may be	available from the view snapshot screen	
The live webpage may provide more currer View catalogue record in Trove	it information	
Copyright information		



Synthetics



Sites that change over time (2024)

Filter Advanced Search

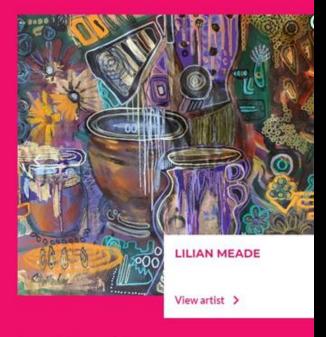


ABOUT HELP NEWS PART

RIVER OF ART FESTIVAL

Blend arts & awe every Spring in the beating heart of the South Coast – Eurobodalla! Discover more than 100 events over 10 days & 100 kilometres of inspiring coastline. The River of Art Festival will be back in 2024. 20-29 SEPTEMBER 2024

Subscribe for Festival Updates >



Featured Art Directory Artists



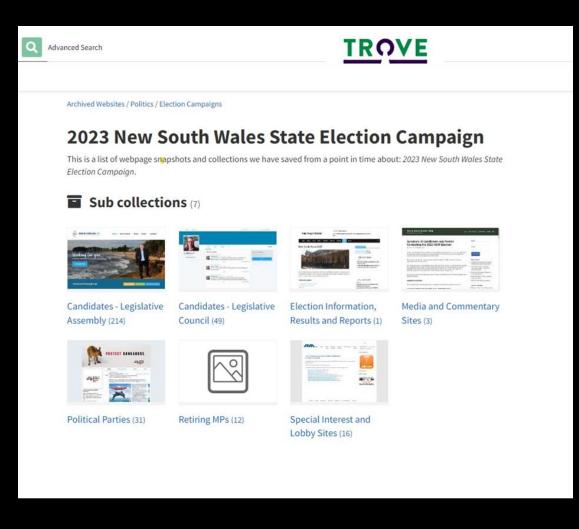
Eurobodalla River of Art Festival (2023)

Filter 👻 🔍 Advanced Search	TROVE	ABOUT HELP NEW!		
RIVER OF ART FESTIVAL				
		Join us again 15 - 24 September 2023 Thanks for making the 2022 Festival the "best ever". The 2022 River of Art Festival celebrated a record 130 arts events across the NSW South Coast. Over 420 artists and 12,500 attendees joined in with music and performances, workshops, open studios, exhibitions, Art on Parade and the River of Art Prize. We closed the Festival with our feature event, Luminous: Art After Dark, a sensational laser and light show on the Moruya River that was attended by over 4,000 people. And the one word that describes the experience was HAPPY. Read our Key Outcomes report for a full breakdown.		



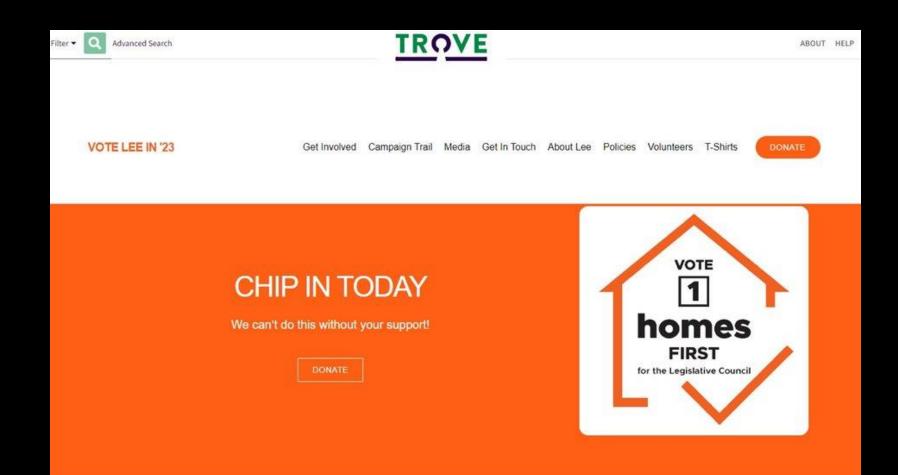
Webpage snapsh

Collections of multiple sites



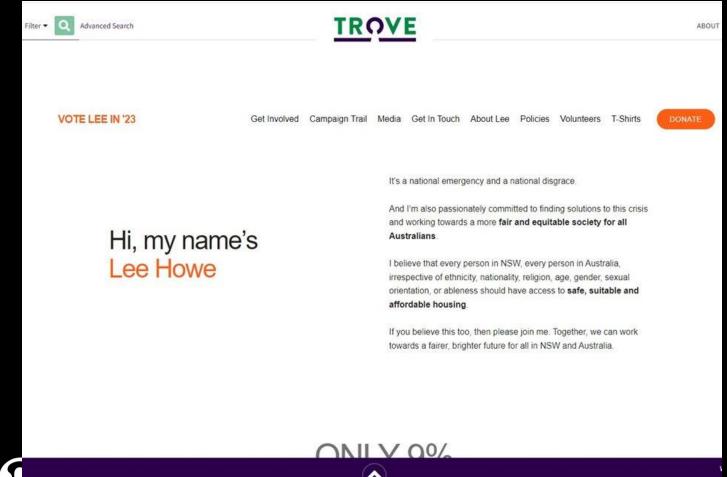


Candidate website / Lee Howe



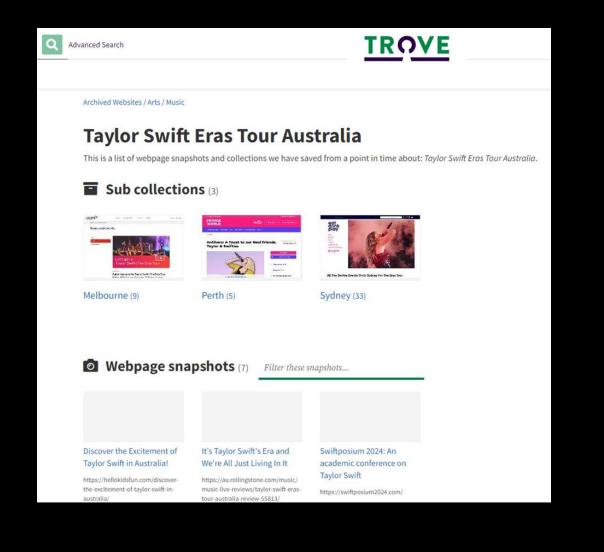


Candidate website / Lee Howe



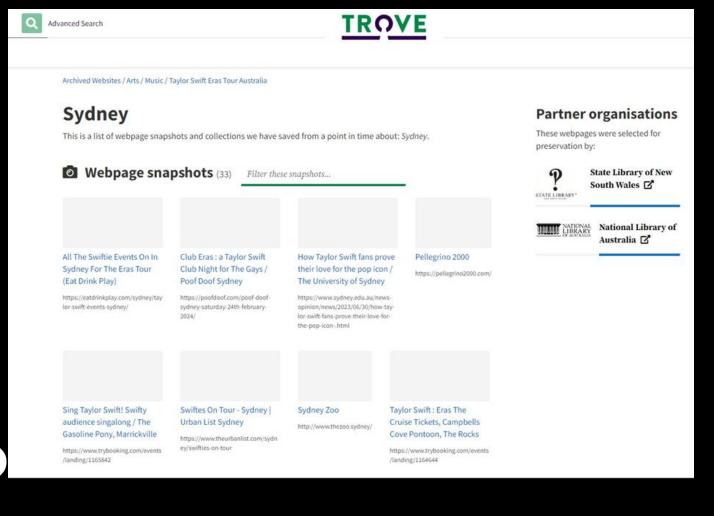


Taylor Swift





Taylor Swift / Sydney

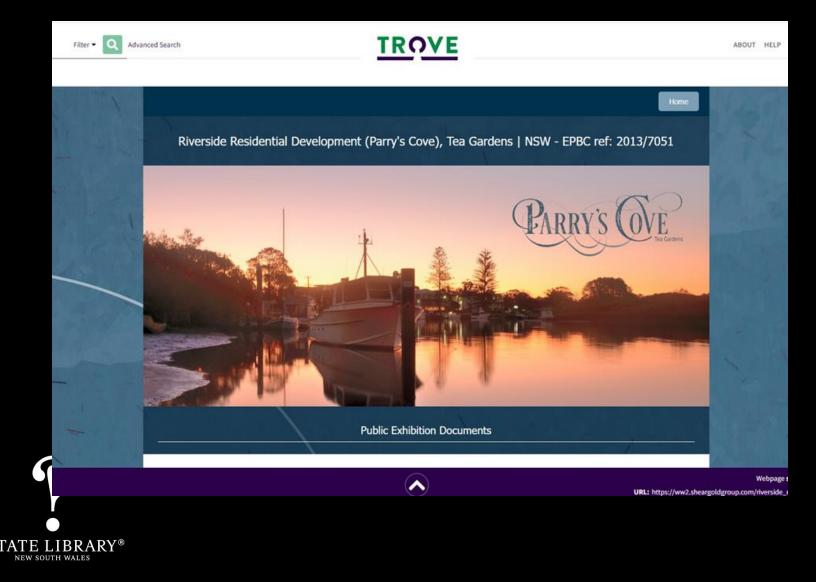


STATE LIBRARY®

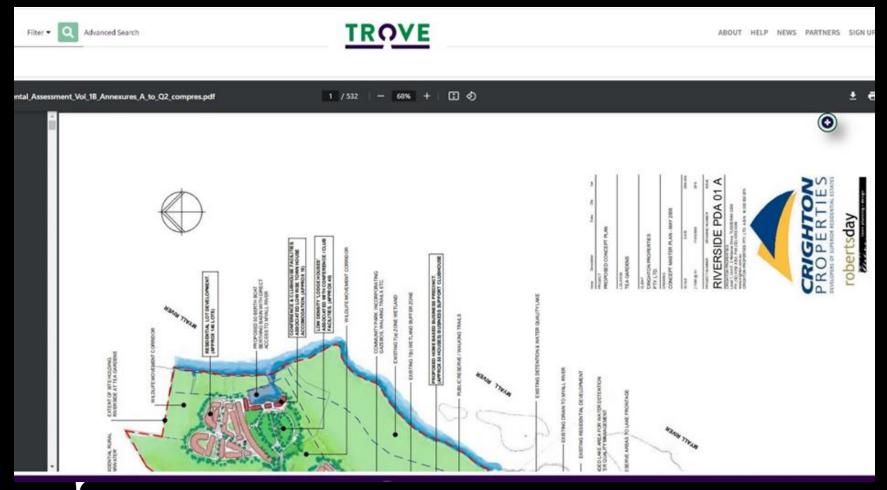
What PANDORA has issues collecting or not suitable for collecting

- 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.
 Paywalled or sites requiring logins.

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



Single PDFs (-> NED)



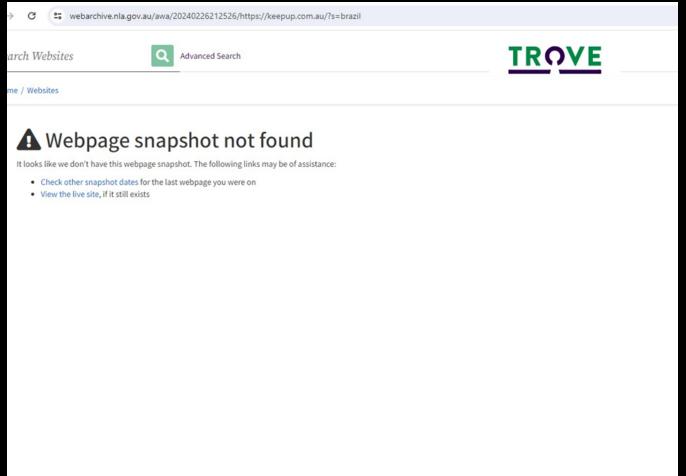
STATE LIBRARY®

Websites with search boxes

Advanced Search	TROVE	ABOUT HELP
KEEPUR LATEST VIDEOS A-LEAGUES AUSSIES ABROAD	MATILDAS SOCCEROOS MORE V FANTASY & TIPPING 7	ALEAGUES.COM AU 7 X MY ACCOM
TRENDING STORIES		
[KEEPUP]	Image: Section of the section of th	head in FA Cup blockbuster
3 WEERS AGO KEEPUP to shut down on March 1: Thank you and goodbye	I MONTH AGO + WOMEN'S SUPER LI Sam Kerr earns second FIF. FIFPro World 11 selection: 'Hugely special'	
	\diamond	Webpage + URL: https://keepup.com.au/

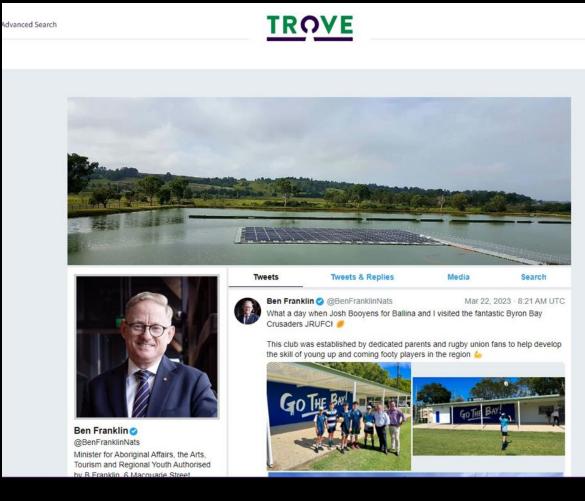
STATE LIBRARY®

KeepUp





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





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



PANDORA – How to find in Trove

- Search with URL
- Search with keywords



With URL

https://cecopaint.wixsite.com/ranca-nsw		Q
		Advanced Search 💌
NEWSPAPERS & GAZETTES	No results	
MAGAZINES & NEWSLETTERS	No results	
IMAGES, MAPS & ARTEFACTS	No results	
RESEARCH & REPORTS	No results	
BOOKS & LIBRARIES	No results	
DIARIES, LETTERS & ARCHIVES	No results	
MUSIC, AUDIO & VIDEO	No results	
PEOPLE & ORGANISATIONS	No results	
WEBSITES	Click the search button below to load results	

Search websites for https://cecopaint.wixsite.com/ranca-nsw



With keyword

JMLA-Projects HTML snapshot 📋 Thursday 16 February 2017 http://jmla.com.au/proj-ranca.html 6 results from jmla.com.au " Raft It " HTML snapshot 📋 Sunday 04 March 2001 ... The River The Map The Photos The Contact Home What Is Ranca Adventure Club? RANCA Adventure club ... and canyon guides. So as you can see RANCA caters for a very broad activity base. What Do We Do? When its all said ... http://www.raftit.com.au/ranca.html 16 results from raftit.com.au https://cecopaint.wixsite.com/ranca-nsw/hmas-wollongong

HTML snapshot 📋 Tuesday 28 February 2023

https://cecopaint.wixsite.com/ranca-nsw/hmas-wollongong

74 results from wixsite.com

Internation Racing Tours - Ambassador Travel Services

HTML snapshot 📋 Monday 31 October 2022

https://ambassadortravel.com.au/general-non-tour-photos?itemId=exu26n8xpn0kbp85qhryf8qu72sm9f-nx3fw-8wksl-2axnn-d2wfb-ranca-pt7bg

3 results from ambassadortravel.com.au

STATE LIBRARY® NEW SOUTH WALES

Entree Kibbles: Up Close with the Deer @ Kampung Cai Ranca Upas [Ciwidey,

PANDORA – How to find in SLNSW Catalogue

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



SLNSW search

RY	Catalogue	Digital co	ollections	
FEEDBACK				
All Fields V https://cecopaint.wix				CLEAR X
<pre> previous Next > </pre>		SORT BY 🗸	20 PER PA 🗸	Refine your searc
O records selected		1 results	SAVE THIS SEARCH +	Online
JOURNALS RANCA NSW / Erica Walker Author / Creator: Walker, Erica Publisher: [Sydney, N.S.W.] : RJ	a (Secretary of RANCA NSW))	:	FORMAT
				SUBJECT
				LANGUAGE
				DATES
LIBRARY [®]				

SLNSW catalogue record

	Alternative Titles	< Prev Bac
	Full title	
	RANCA NSW / Erica Walker.	÷ ⇔
	Authors, Artists and Contributors	
	Author / Creator	How to access t
	Walker, Erica (Secretary of RANCA NSW)	Help Ask a Libr
	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	
	Archived version	
IDENTIFIERS	Primary Identifiers	
	Record Identifier	
	74VKabryZ7KZ	
	Permalink	
	https://collection.sl.nsw.gov.au/record/74VKabryZ7KZ	
	COPY LINK 8	
	Other Identifiers	
	DDC	
	359.3254	
	MMS ID	



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



