

From: [REDACTED]
Sent: Thursday, 18 August 2011 3:13 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: making a decision about the Argus Online project

TRIMmed
DATE:
No: R11/80720
P3

I believe we should go with Option 1 - Trove.

There would be a significant amount of work to get the data into the old Argus Index (option 2) given the amount of time that has passed since we last loaded data into it. If we're going to do work anyway, it may as well go into Trove. I don't think it would be difficult but it could still be fairly time-consuming - last time there were quite a few structural mistakes in the data that we had to return to [REDACTED] to fix. And there would have to be some changes to Trove.

[REDACTED]

From: [REDACTED]
Sent: Thursday, 18 August 2011 2:50 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: making a decision about the Argus Online project

[REDACTED]
Following on from the recent Trove/ANDP team discussion about continuing the Argus Online project, I would welcome feedback on the following (by Friday 26 August if possible). Of especial significance to help us make a final decision will be the amount of time required to deliver each option.

[REDACTED] has provided me with some usage statistics for the site, which I requested to determine the essentiality of the service. Deirdre gave me some comparative statistics for other products which the Library pays for.

Argus Online
2007 - 13,645 (8.48 pages per visit)
2008 - 19,192 (7.69 pages per visit)
2009 - 17,988 (7.88 pages per visit)
2010 - 21,231 (9.11 pages per visit)
2011 (to date) - 7,922 (6.82 pages per visit)

As previously mentioned, I believe the 2010 increase is due to some publicity about the impending index updates [<http://blogs.abc.net.au/victoria/2010/05/the-argus-makes-a-return-online.html>]. However, these statistics indicate a reasonable level of interest, the availability of the digitised pages not withstanding.

eResources access

	To date
Onsite and offsite	this year

Subtotal via catalogue	11,406
Subtotal via eResources	52,501
Subtotal via Trove	10
Total	63,917

Options for proceeding

Option 1

In line with our strategic objective of reducing the number of separate websites the Library supports, we could treat the indexes as a source of data (like APAIS data) for Trove. Responsibilities to do this are suggested as follows:

CIB and CAB

- With assistance from [REDACTED], who is familiar with the data, confirm that the next set of index data can be mapped and loaded into Trove. Instead of being displayed as a separate service, integrate the display into Trove (could this be done in a way similar to the labelling being proposed for the AWW?). Could it also form a test case for other indexes?
- On load, the index entries should link to the relevant pages in Trove. A better level of granularity is unlikely because the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.

Trove support team

- With relevant IT staff, contact the index creators and propose the new approach

Web Publishing Branch

- Subject to confirmation, eventually switch off the old website.

Option 2

If it is not possible to load the data into Trove, then the likely actions are:

Web Publishing Branch

- Change the current website to reflect the availability of a new decade
- Provide a link to Trove to indicate the availability of fulltext.

CAB and CIB

- Load the new indexes using existing software, with assistance from [REDACTED]

Trove support team

- Advise the index creators that data will be loaded only for the remainder of the decade, not to 2003 as proposed in the original grant.

Option 1 is preferred because:

- it reduces the number of siloed websites hosted by the NLA
- it adds value to Trove
- it satisfies the request to be a part of Trove
- new data can be added using a routine which does not have to be dusted off every few years or so.

Thanks for your advice,

[REDACTED]

[REDACTED]
[REDACTED]
Resource Sharing and Innovation Division
National Library of Australia

Parkes Place
Canberra ACT 2601

e: [REDACTED]

TRIMMED
DATE:

NO: R11/50718
PS

From: [REDACTED]
Sent: Friday, 19 August 2011 11:51 AM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: making a decision about the Argus Online project

Hi,

I think [REDACTED] suggestion of loading the index as a separate resource rather than only mapping the index to articles in Trove is good because:

- It will often be hard to correctly identify the article. I'm not sure of the extent of havoc caused by the digitisation of a different edition (in newspapers) from the one indexed in Argus Index ([REDACTED] sample suggests around 20% inaccuracies just on page number), but regardless, the article zoning in newspapers is often coarse and the Argus Index does not identify articles, only date, page and column
- This approach "honours the intent" of [REDACTED] and her indexers, who put a lot of store in the structure of the index, and were even vocally opposed to searching (as opposed to browsing) of the index in the initial version. The custom Trove resource display for each heading would probably look a lot like the current Argus Index entry for a heading.
- Many of the Argus Index headings and descriptions are hard to "translate" to tags, but that is currently the only mechanism we have for representing them. They are not comments. For example, typical headings are
 - o Abattoirs -
 - Henry Walker, boiling down works nuisance
 - overflow at Richmond abattoirs after storm
 - St Kilda abattoirs
 - inspector of nuisances at Hawthorn reports on inaction on paving and draining of butchers yards
 - Central Board of Health objections to slaughtering of animals on butcher's premises, Fitzroy

Many are extremely long.

But:

- Although this information is invaluable (not least because it provides a entry to articles whose OCR quality means the current NDP text doesn't), if it is kept separate from newspapers (in another zone), we'd have to make sure newspaper searchers were drawn to its existence – it would be disappointing if it were only visible from a non-newspaper island. So I think we have to incorporate it somehow in the NDP index as well (more below)
- How will indices from multiple years be combined or represented? Maybe this is a non-issue because the indexers have been consistent and not changed their minds on index structure as they've been going (or maybe they've retrospectively amended previous indices, and we need to acquire those amendments) – maybe [REDACTED] knows this!?
- Will we be able to hazard a guess at the matching article in NDP? For example, if you came across this article

[REDACTED]

it would be nice to see the index entry: "Extradition - Sergeant Toohey, from NZ to arrest John Fishenden, arrested himself for leaving with money Fishenden procured by fraud" and particularly to know the court case had been indexed **with the same heading** "3 Jul 1878 5d [court case]" and link straight to that page/column: <http://trove/ndp/del/article/5938586/250597?searchTerm=toohey#pstart250597>

These are both long, multipage (but typical) articles, which demonstrate the problem of precise assignment of the indexing and finding or navigation to the "right" part, but perhaps by adding navigation to a column (rather than just to article or page) can help matters.

- How would we show the index contents in the newspapers display? Will we index them with the article (if we can guess it), or articles on that page/column if we cant? Will we show them in newspaper search results? (As snippets or as we currently show tags?) Will we show them on the page and/or article detail displays similar to how we show tags, or will we somehow associate them with column headings? Will we display them in full or hint at them and reveal them when user's curiosity is piqued? (I think we just have to try a few reasonable things and test them with real users.)
- How will we minimise the confusion caused by the distinction between tags and Argus Index descriptors? Will users be puzzled that index only covers Argus, and will they assume it covers more titles and worry about omissions?
- Will Argus Index entries be correctable/deletable/addable? For example, in the above extradition story, MAYBE the Argus Indexers missed an article. It would also be tempting to extend the index to other titles (perhaps some other title contained a follow-up or different interpretation of this court case), but how would this interact with tagging – would it create more confusion than the separate Argus Index data will inevitably cause?

Another reason for "correction" would be page and column errors (see [redacted] Word document), but then are we changing the Argus Index to refer to the Sydney edition, or would we retain the Melbourne edition reference and create a separate one (which NDP uses) for the Sydney version?

--

Re the surge in Argus Online usage 2010, 7 days in May accounted for over 80% of that month's traffic, which itself was about 6 times more than the monthly average, and most of the May traffic came from a Russian IP address which looks to be a crawler masquerading as a browser. The page view average recorded by AWW stats for July 2011 is 175, a stat that doesn't seem to have obvious inaccuracies caused by misidentified robots or bogus page counting. But the stats are very volatile – there are numerous flurries of activity (eg 28Nov10 – 5K pages, 1 Nov10 – 2.5K pages, 12Aug10 – 8K pages) – I've looked at a few of these and they typically look like a site crawl from a single IP address (not identified as a robot by AWW stats).

By way of comparison, Argus pages on Trove typically get over 100K hits per day, but those Argus pages cover a much longer period of time (1848 – 1956) and are of course full text!

[redacted]

From: [redacted]
Sent: Friday, 19 August 2011 11:16 AM
To: [redacted]
Subject: RE: making a decision about the Argus Online project

Hi,

Here's some information to start. A guess at accuracy of page numbers, and how many keywords from the index are searchable in Trove.

<< File: Argus.docx >>

From: [REDACTED]
Sent: Thursday, 18 August 2011 3:10 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: FW: making a decision about the Argus Online project

Hi [REDACTED]

Here is another piece of work for you. Background emails to follow.

Could you investigate please ([REDACTED] and [REDACTED] worked on the Argus index in the past) what the options are, issues would be, and work involved in adding this content into Trove. (A standout issue I think is that the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.)

Is a timeframe of 26 August achievable for you to do this? If not, when would be?

From: [REDACTED]
Sent: Thursday, 18 August 2011 2:50 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: making a decision about the Argus Online project

[REDACTED]

Following on from the recent Trove/ANDP team discussion about continuing the Argus Online project, I would welcome feedback on the following (by Friday 26 August if possible). Of especial significance to help us make a final decision will be the amount of time required to deliver each option.

[REDACTED] has provided me with some usage statistics for the site, which I requested to determine the essentiality of the service. [REDACTED] gave me some comparative statistics for other products which the Library pays for.

Argus Online

2007 – 13,645 (8.48 pages per visit)
2008 – 19,192 (7.69 pages per visit)
2009 – 17,988 (7.88 pages per visit)
2010 – 21,231 (9.11 pages per visit)
2011 (to date) - 7,922 (6.82 pages per visit)

As previously mentioned, I believe the 2010 increase is due to some publicity about the impending index updates [<http://blogs.abc.net.au/victoria/2010/05/the-argus-makes-a-return-online.html>]. However, these statistics indicate a reasonable level of interest, the availability of the digitised pages notwithstanding.

eResources access

	To date this year
Onsite and offsite	
Subtotal via catalogue	11,406
Subtotal via eResources	52,501
Subtotal via Trove	10
Total	63,917

Options for proceeding

Option 1

In line with our strategic objective of reducing the number of separate websites the Library supports, we could treat the indexes as a source of data (like APAIS data) for Trove. Responsibilities to do this are suggested as follows:

CIB and CAB

- With assistance from [REDACTED], who is familiar with the data, confirm that the next set of index data can be mapped and loaded into Trove. Instead of being displayed as a separate service, integrate the display into Trove (could this be done in a way similar to the labelling being proposed for the AWW?). Could it also form a test case for other indexes?
- On load, the index entries should link to the relevant pages in Trove. A better level of granularity is unlikely because the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.

Trove support team

- With relevant IT staff, contact the index creators and propose the new approach

Web Publishing Branch

- Subject to confirmation, eventually switch off the old website.

Option 2

If it is not possible to load the data into Trove, then the likely actions are:

Web Publishing Branch

- Change the current website to reflect the availability of a new decade
- Provide a link to Trove to indicate the availability of fulltext.

CAB and CIB

- Load the new indexes using existing software, with assistance from [REDACTED]

Trove support team

- Advise the index creators that data will be loaded only for the remainder of the decade, not to 2003 as proposed in the original grant.

Option 1 is preferred because:

- (a) it reduces the number of siloed websites hosted by the NLA
- (b) it adds value to Trove
- (c) it satisfies the request to be a part of Trove
- (d) new data can be added using a routine which does not have to be dusted off every few years or so.

Thanks for your advice,

[REDACTED]

[REDACTED]

[REDACTED]
Resource Sharing and Innovation Division
National Library of Australia

Parkes Place
Canberra ACT 2601

e: [REDACTED]

Jan Harris

TRIMMED

From: [REDACTED]
Sent: Friday, 19 August 2011 11:16 AM
To: [REDACTED]
Subject: RE: making a decision about the Argus Online project

1601 B11/80719
P8

Hi,

Here's some information to start. A guess at accuracy of page numbers, and how many keywords from the index are searchable in Trove.



Argus.docx

[REDACTED]

From: [REDACTED]
Sent: Thursday, 18 August 2011 3:10 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: FW: making a decision about the Argus Online project

Hi [REDACTED]

Here is another piece of work for you. Background emails to follow.

Could you investigate please ([REDACTED] and [REDACTED] worked on the Argus index in the past) what the options are, issues would be, and work involved in adding this content into Trove. (A standout issue I think is that the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.)

Is a timeframe of 26 August achievable for you to do this? If not, when would be?

[REDACTED]

From: [REDACTED]
Sent: Thursday, 18 August 2011 2:50 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: making a decision about the Argus Online project

[REDACTED]

Following on from the recent Trove/ANDP team discussion about continuing the Argus Online project, I would welcome feedback on the following (by Friday 26 August if possible). Of especial significance to help us make a final decision will be the amount of time required to deliver each option.

██████ has provided me with some usage statistics for the site, which I requested to determine the essentiality of the service. Deirdre gave me some comparative statistics for other products which the Library pays for.

Argus Online

2007 – 13,645 (8.48 pages per visit)
2008 – 19,192 (7.69 pages per visit)
2009 – 17,988 (7.88 pages per visit)
2010 – 21,231 (9.11 pages per visit)
2011 (to date) - 7,922 (6.82 pages per visit)

As previously mentioned, I believe the 2010 increase is due to some publicity about the impending index updates [<http://blogs.abc.net.au/victoria/2010/05/the-argus-makes-a-return-online.html>]. However, these statistics indicate a reasonable level of interest, the availability of the digitised pages notwithstanding.

eResources access

	To date this year
Onsite and offsite	
Subtotal via catalogue	11,406
Subtotal via eResources	52,501
Subtotal via Trove	10
Total	63,917

Options for proceeding

Option 1

In line with our strategic objective of reducing the number of separate websites the Library supports, we could treat the indexes as a source of data (like APAIS data) for Trove. Responsibilities to do this are suggested as follows:

CIB and CAB

- With assistance from ████████, who is familiar with the data, confirm that the next set of index data can be mapped and loaded into Trove. Instead of being displayed as a separate service, integrate the display into Trove (could this be done in a way similar to the labelling being proposed for the AWW?). Could it also form a test case for other indexes?
- On load, the index entries should link to the relevant pages in Trove. A better level of granularity is unlikely because the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.

Trove support team

- With relevant IT staff, contact the index creators and propose the new approach

Web Publishing Branch

- Subject to confirmation, eventually switch off the old website.

Option 2

If it is not possible to load the data into Trove, then the likely actions are:

Web Publishing Branch

- Change the current website to reflect the availability of a new decade
- Provide a link to Trove to indicate the availability of fulltext.

CAB and CIB

- Load the new indexes using existing software, with assistance from [REDACTED]

Trove support team

- Advise the index creators that data will be loaded only for the remainder of the decade, not to 2003 as proposed in the original grant.

Option 1 is preferred because:

- (a) it reduces the number of siloed websites hosted by the NLA
- (b) it adds value to Trove
- (c) it satisfies the request to be a part of Trove
- (d) new data can be added using a routine which does not have to be dusted off every few years or so.

Thanks for your advice,

[REDACTED]

[REDACTED]

[REDACTED]

Resource Sharing and Innovation Division
National Library of Australia

Parkes Place
Canberra ACT 2601

e: [REDACTED]

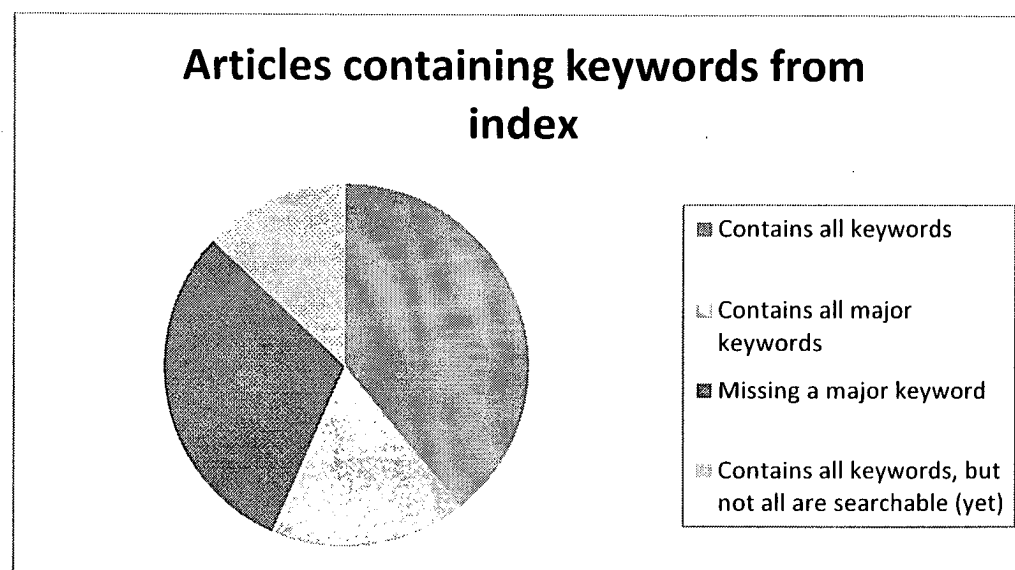
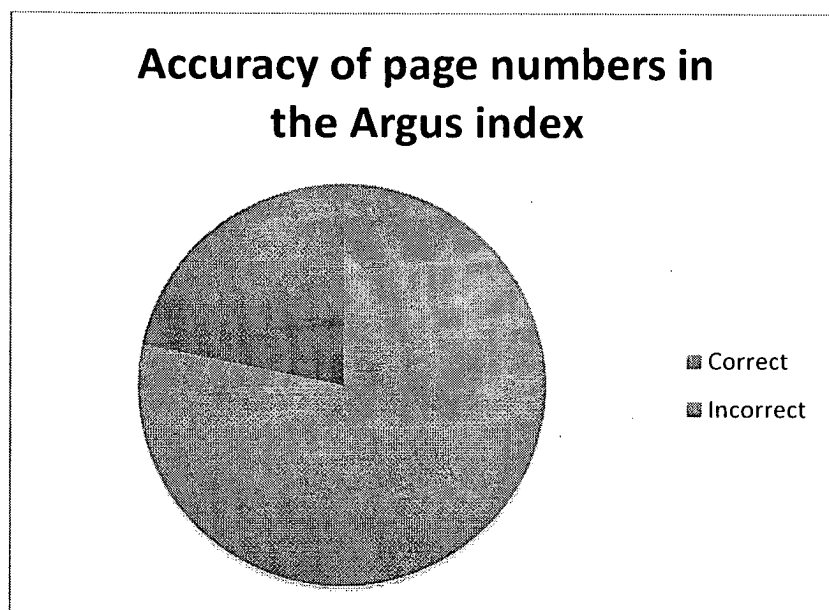
Analysis of Argus Index Data

Overview

23 articles listed under three different headings in the Argus index were reviewed.

- 4 articles about “Aarons, Joseph” from 1870-1
- 11 articles about “Pacific Islands” from 1875
- 8 articles about “Waygunyah” from 1879

From this sample, the page number correlated between the index and the digitised article online. About one third of online articles were missing a major keyword from the index. All three indexes were missing some relevant articles found by searching online.



Examples of missing minor keywords:

- Creation
- Capt (article contains “captain” instead)
- Complain (article contains complaint instead)

Aarons, Joseph

<http://www.nla.gov.au/apps/argus?action=BrowseSingleHeading&showOnly=AARONS,%20JOSEPH&year=1870&suppressPseud=n>

Argus Index				Trove				
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords
complains of nonreceipt of funds raised for masonic charities	1 Feb 1870	5	b	http://trove.nla.gov.au/ndp/del/page/225249	4	f	N	complain (complaint) nonreceipt (receipt)
election for Latrobe Ward	10 Feb 1870	5	a	http://trove.nla.gov.au/ndp/del/page/225380	4	d	Y	
election for Latrobe Ward	11 Feb 1870	5	c	http://trove.nla.gov.au/ndp/del/page/225396	4	e	Y	
Haydamack lunch	3 Apr 1871	6	f[letter]	http://trove.nla.gov.au/ndp/del/page/229661	6	f	Y	

Some articles found online are missing from this index:

Election notices: <http://trove.nla.gov.au/ndp/del/article/5812153?searchTerm=Aarons%20joseph> "*ELECTORB of LATROBE WARD, Allow mo to oxproBs my sincere acknowledge- ment for the honour conferred In olcctlng mo your representativo in the City Council. ' Whilst I am permitted to onjoy thit privilege no exertion will bo wanting to further tho interests of your ward and the city. Yours most obedlontly, JOSEPH AARONS.*"

Pacific Islands

<http://www.nla.gov.au/apps/argus?action=BrowseIndex&year=1875&startFrom=p>

Argus Index				Trove				
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords
Capt. Macauley, Margaret and Jane, assaults female islander	2 Jan 1875	5	d	http://trove.nla.gov.au/ndp/del/page/239310	5	d	N	capt (captain)
King Kalakana, Sandwich Islands, visits Vic.	10 Feb 1875	10	d	http://trove.nla.gov.au/ndp/del/page/239908	10	d	Y	
Commodore J. G. Goodenough murdered by natives of Santa Cruz	24 Aug 1875	4	f [editorial]	http://trove.nla.gov.au/ndp/del/page/242956	4	g	N	pacific islands
“	24 Aug 1875	4	g	http://trove.nla.gov.au/ndp/del/page/242956	4	g	N	pacific islands
“	24 Aug 1875	5	g	http://trove.nla.gov.au/ndp/del/page/242957	5	g	N	Murdered J. G.
“	24 Aug 1875	6	a	http://trove.nla.gov.au/ndp/del/page/242959	6	a	N	J. G. (Jas. Graham) Pacific
“	28 Aug 1875	7	e	http://trove.nla.gov.au/ndp/del/page/243025	7	g	N	Murdered Natives Santa Cruz
“	30 Aug 1875	7	d	http://trove.nla.gov.au/ndp/del/page/243049	7	d	N	Murdered Natives Santa Cruz J. G. (James Graham)
“	8 Oct 1875	5	g	http://trove.nla.gov.au/ndp/del/page/241476	5	g	N	Murdered Natives Santa Cruz
“	15 Oct 1875	6	d	http://trove.nla.gov.au/ndp/del/page/241497	5	g	Y	Cruz (Cru,,)
“	16 Oct 1875	4	c	http://trove.nla.gov.au/ndp/del/page/241604	4	c	Y	

Some articles found online are missing from this index:

31 August 1875 7a <http://trove.nla.gov.au/ndp/del/article/11522960> *"The death of Commodore Goodenough of which you have already received the full particulars has caused unfeigned sorrow in Sydney..."*

6 September 1875 1Sa <http://trove.nla.gov.au/ndp/del/article/7419003> *"H.M.S. Pearl arrived in Port Jackson, with the dead bodies of Commodore Goodenough and two seamen belonging to his vessel on board, all three having fallen victims to the treachery of the Savages of Santa Cruz..."*

9 Oct 1875 6b <http://trove.nla.gov.au/ndp/del/article/7421860> *"The news of the murder of Commodore Goodenough caused great sorrow and mourning at Levuka. Seldom if ever is the loss of a man so universally felt..."*

Wahgunyah

<http://www.nla.gov.au/apps/argus?action=BrowseIndex&year=1879&startFrom=WAGES&headingType=%20&inc=n>

Argus Index				Trove					
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords	
opening of line from Springs to Wahgunyah	16 Jan 1879	5	a	http://trove.nla.gov.au/ndp/del/page/251508	5	a	Y	Wahgunyah (Wah"vmjah)	
opening of line from Springs to Wahgunyah	27 Jan 1879	5	a	http://trove.nla.gov.au/ndp/del/page/251674	5	a	Y	Wahgunyah (Waygunsah)	
opening of line from Springs to Wahgunyah	30 Jan 1879	5	e	http://trove.nla.gov.au/ndp/del/page/251721	5	e	Y		
creation as licensing district	3 Sep 1879	5	b	http://trove.nla.gov.au/ndp/del/page/252911	5	b	Y		
creation as licensing district	5 Sep 1879	7	c [letter]	http://trove.nla.gov.au/ndp/del/page/252959	7	c	N	creation	
creation as licensing district	9 Sep 1879	7	f [letter]	http://trove.nla.gov.au/ndp/del/page/253010	7	f	N	creation	
pleuro-pneumonia	18 Nov 1879	7	b	http://trove.nla.gov.au/ndp/del/page/254099	7	b	Y		
pleuro-pneumonia	1 Dec 1879	5	b	http://trove.nla.gov.au/ndp/del/page/254287	6	g	Y		

Some articles found online are missing from this index:

6 Feb 1879 7a Injury to Wahgunyah railway <http://trove.nla.gov.au/ndp/del/article/5931330?searchTerm=wahgunyah>

31 Oct 1879 3c Flood-waters at Wahgunyah <http://trove.nla.gov.au/ndp/del/article/5961491?searchTerm=wahgunyah>

29 Oct 1879 Purchase of Wahgunyah bridge <http://trove.nla.gov.au/ndp/del/article/5961166?searchTerm=wahgunyah>

TRIMmed

From: [REDACTED]
Sent: Tuesday, 23 August 2011 4:15 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: making a decision about the Argus Online project

DATE:

No: R11/80215
Pg 8

Hi,

I think tags are a good option because they already do a lot of what we want (add extra search terms and phrases to the article, link related articles) without needing to add anything extra to the interface or back-end.

Speaking to [REDACTED], he remembers quickly trialling a way to identify articles once before, and thought it was quite successful, so that is sounding possible after all.

"I tried searching the issue using words from the argus heading and sub heading after removing common words, getting article with most and densest hits"

He also suggested another option would be to do an entirely different user interface for the indexes (similar to the one in the pipeline for AWW). For example, with a date slider etc.

I think we could also do a combination of 1 & 2 or 1 & 3, if we liked.

I don't think it matters much to us how long the project takes to produce their indexes, does it? But I agree, to proceed down the current path seems like madness. Maybe we could just ask [REDACTED] team to add tags to the articles in Trove in future?

[REDACTED], do you have enough information to do some estimates for the four different proposed solutions?

1. Add a record to Trove for each major heading (and display within Trove)
2. Add index annotations to the digitised article in Trove
3. Add index data to the digitised article in Trove, as tags
4. In addition to solution 1, add a specialised interface for viewing the index

[REDACTED] estimates the AWW interface one will take 1.5-2 months of design and implementation work)

From: [REDACTED]
Sent: Tuesday, 23 August 2011 3:35 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: making a decision about the Argus Online project

My 2c worth:

To be frank, the usefulness of the Argus Index in its previous form has been superseded by Trove newspapers.

For that reason putting effort into [REDACTED] option 1 is less desirable but it does avoid the problem of identifying articles because you could just link to the page.

The second and third options rely on automatically identifying the articles, perhaps by trying to match keywords, date and page number and for those that don't match, looking for articles that start in the area suggested by the column letter. We'd have to do a trial of identifying articles to see if this was reliable enough to be worth doing. The tag option seems a good option.

If we can reliably identify articles we'd have to discuss with [redacted] to see if this option was acceptable, since it doesn't preserve the integrity (browsability) of the index. If it is unacceptable (or if it turns out we can't reliably identify articles) the options would be a) go with option 1, b) add the new data to the existing Argus Index website or c) maybe even tell her we can't accept new data.

If we did go with the tag option and [redacted] wanted to keep on indexing in future, a condition of undertaking to accept the data should be that she indexes using Trove newspapers rather than print copies and cites the article ids instead of just page and column numbers. But personally I wouldn't encourage her to keep indexing - I think we should make it clear we can't commit to accepting data in future especially [redacted]

[redacted] I'm not even sure we should put much effort into investigating any option, other than identifying articles, [redacted]

From: [redacted]
Sent: Tuesday, 23 August 2011 11:54 AM
To: [redacted]
Cc: [redacted]
Subject: RE: making a decision about the Argus Online project

Hi Everyone,

Some responses in situ below, in red.

Trial of adding tags:

Investigating the tag idea further, I added one of the longer abattoir index terms as a tag to this article:

<http://trove.nla.gov.au/ndp/del/article/5813133?searchTerm=abattoirs>

The current limit is 60 characters for a person adding tags, so I had to break it up into two tags, but we could ignore that limit when automatically adding this data. The good thing about tags is that if you click on them you see all the other items with this tag, and this is already built in. I think our current display doesn't make it easy to see where one tag ends and the next starts, but we could easily fix that. I don't think this solution is too bad.

I also tried adding an index term as a comment, which isn't as useful because you can't click to see other items with the same comment. Since the index terms are a description of the article content, I think they fit better conceptually into the model of tags, rather than comments, anyway.

<http://trove.nla.gov.au/ndp/del/article/5831982?searchTerm=slaughter>

Summary of proposed solutions:

Feature	Benefits	Challenges
Add a record to Trove for each major heading (e.g. Abattoirs)	<ul style="list-style-type: none">• Doesn't add clutter to NDP view• Gets the data into Trove and searchable• Preserves the structure of the	<ul style="list-style-type: none">• Will newspaper users find it?• Is there something we can link to from the Argus title page, or elsewhere?• In isolation, doesn't allow people

	<p>index: the sub-headings and the curated list of articles on a topic</p>	<p>looking at an article see the index/related articles</p> <ul style="list-style-type: none"> • A particular heading can still be browsed, but what about a year? • Which zone? • How to link to the right article? • Requires a custom display to be written • Does this lend itself to user correction, or not?
<p>Add index annotations (not tags or comments) to the digitised article</p>	<ul style="list-style-type: none"> • Immediately accessible to newspapers users • Could link to other articles with the same index term (reminiscent of tags) • Could link to the next article with the same index term, if useful?(chronologically) 	<ul style="list-style-type: none"> • How to preserve sub-heading structure? • In isolation, might not preserve the structure of the index – no overview or ability to browse? • How to avoid clutter/confusion on the article screen? • Does this lend itself to user correction, or not? • Requires a custom display to be written • How to identify the right article?
<p>Add index data to the digitised article as tags</p>	<ul style="list-style-type: none"> • Immediately accessible to newspaper users • Doesn't add new clutter to article screen – uses existing page sections • Automatically provides links to other articles with the same tag/index term 	<ul style="list-style-type: none"> • How to identify the right article? • Can't preserve sub-heading structure • In isolation, might not preserve the structure of the index – no overview or ability to browse? • Does this lend itself to user correction, or not? • Too long for existing tag structure, or can we work around this? (limit 60)

If we decide to "correct" the index to match the digitised version of the paper, will we ask [redacted] to start indexing the Sydney version of the paper in future instead?

See below for responses to Kent, in red, including a mockup of an Argus index record in Trove.

[redacted]

From: [redacted]
Sent: Friday, 19 August 2011 11:51 AM
To: [redacted]
Cc: [redacted]
Subject: RE: making a decision about the Argus Online project

Hi,

I think [redacted] suggestion of loading the index as a separate resource rather than only mapping the index to articles in Trove is good because:

- It will often be hard to correctly identify the article. I'm not sure of the extent of havoc caused by the digitisation of a different edition (in newspapers) from the one indexed in Argus Index ([redacted] sample suggests around 20% inaccuracies just on page number), but regardless, the article zoning in newspapers is often coarse and the Argus Index does not identify articles, only date, page and column

[redacted]: I think if we can't find a way to reliably identify the article, it will limit what we can usefully do.

- This approach "honours the intent" of [redacted] and her indexers, who put a lot of store in the structure of the index, and were even vocally opposed to searching (as opposed to browsing) of the index in the initial version. The custom Trove resource display for each heading would probably look a lot like the current Argus Index entry for a heading.

[redacted]: Would this content be "hosted" in Trove, Newspapers, or elsewhere? Is it a static webpage or a record ingested by Trove or both? How do we classify the format – "other article"? I'm not sure what people have in mind. Or what is sensible.

See attached mockup for an idea on how this might look. I don't think this mockup really explains to people what this is.

<< File: screenshot_argusindex copy.jpg >>

- Many of the Argus Index headings and descriptions are hard to "translate" to tags, but that is currently the only mechanism we have for representing them. They are not comments. For example, typical headings are
 - o Abattoirs -
 - Henry Walker, boiling down works nuisance
 - overflow at Richmond abattoirs after storm
 - St Kilda abattoirs
 - inspector of nuisances at Hawthorn reports on inaction on paving and draining of butchers yards
 - Central Board of Health objections to slaughtering of animals on butcher's premises, Fitzroy

Many are extremely long.

But:

- Although this information is invaluable (not least because it provides a entry to articles whose OCR quality means the current NDP text doesn't), if it is kept separate from newspapers (in another zone), we'd have to make sure newspaper searchers were drawn to its existence – it would be disappointing if it were only visible from a non-newspaper island. So I think we have to incorporate it somehow in the NDP index as well (more below)
- How will indices from multiple years be combined or represented? Maybe this is a non-issue because the indexers have been consistent and not changed their minds on index structure as they've been going (or maybe they've retrospectively amended previous indices, and we need to acquire those amendments) – maybe [redacted] knows this!?
- Will we be able to hazard a guess at the matching article in NDP? For example, if you came across this article

<http://trove/ndp/del/article/5937959/250524?searchTerm=toohey#pstart250524>

it would be nice to see the index entry: "Extradition - Sergeant Toohey, from NZ to arrest John Fishenden, arrested himself for leaving with money Fishenden procured by fraud" and particularly to know the court case had been indexed **with the same heading** "3 Jul 1878 5d [court case]" and link straight to that page/column: <http://trove/ndp/del/article/5938586/250597?searchTerm=toohey#pstart250597>

These are both long, multipage (but typical) articles, which demonstrate the problem of precise assignment of the indexing and finding or navigation to the "right" part, but perhaps by adding navigation to a column (rather than just to article or page) can help matters.

█: I don't think we can attach anything to a particular article, if we base our guess solely on page number and column - it will be wrong too often. Can we come up with a way to more accurately guess the article? Maybe searching the page and surrounding two(?) pages for the index keywords and selecting the best match? (they regularly aren't all present) Additionally, I don't think the Argus index includes advertising, so we could immediately discount those articles, if that is correct. Does that sound feasible, or not?

- How would we show the index contents in the newspapers display? Will we index them with the article (if we can guess it), or articles on that page/column if we cant? Will we show them in newspaper search results? (As snippets or as we currently show tags?) Will we show them on the page and/or article detail displays similar to how we show tags, or will we somehow associate them with column headings? Will we display them in full or hint at them and reveal them when user's curiosity is piqued? (I think we just have to try a few reasonable things and test them with real users.)

█: We can only do this if we can fairly accurately identify the right article, otherwise I think it is too confusing and misleading.

- How will we minimise the confusion caused by the distinction between tags and Argus Index descriptors? Will users be puzzled that index only covers Argus, and will they assume it covers more titles and worry about omissions?
- Will Argus Index entries be correctable/deletable/addable? For example, in the above extradition story, MAYBE the Argus Indexers missed an article. It would also be tempting to extend the index to other titles (perhaps some other title contained a follow-up or different interpretation of this court case), but how would this interact with tagging – would it create more confusion than the separate Argus Index data will inevitably cause?

█: I would think not editable, at least to start.

Another reason for "correction" would be page and column errors (see █ Word document), but then are we changing the Argus Index to refer to the Sydney edition, or would we retain the Melbourne edition reference and create a separate one (which NDP uses) for the Sydney verion?

--

Re the surge in Argus Online usage 2010, 7 days in May accounted for over 80% of that month's traffic, which itself was about 6 times more than the monthly average, and most of the May traffic came from a Russian IP address which looks to be a crawler masquerading as a browser. The page view average recorded by AWW stats for July 2011 is 175, a stat that doesn't seem to have obvious inaccuracies caused by misidentified robots or bogus page counting. But the stats are very volatile – there are numerous flurries of activity (eg 28Nov10 – 5K pages, 1 Nov10 – 2.5K pages, 12Aug10 – 8K pages) – I've looked at a few of these and they typically look like a site crawl from a single IP address (not identified as a robot by AWW stats).

By way of comparison, Argus pages on Trove typically get over 100K hits per day, but those Argus pages cover a much longer period of time (1848 – 1956) and are of course full text!

█

From: █
Sent: Friday, 19 August 2011 11:16 AM
To: █
Subject: RE: making a decision about the Argus Online project

Hi,

Here's some information to start. A guess at accuracy of page numbers, and how many keywords from the index are searchable in Trove.

<< File: Argus.docx >>

From: [REDACTED]
Sent: Thursday, 18 August 2011 3:10 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: FW: making a decision about the Argus Online project

Hi [REDACTED]

Here is another piece of work for you. Background emails to follow.

Could you investigate please ([REDACTED] and [REDACTED] worked on the Argus index in the past) what the options are, issues would be, and work involved in adding this content into Trove. (A standout issue I think is that the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.)

Is a timeframe of 26 August achievable for you to do this? If not, when would be?

From: [REDACTED]
Sent: Thursday, 18 August 2011 2:50 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: making a decision about the Argus Online project

[REDACTED]

Following on from the recent Trove/ANDP team discussion about continuing the Argus Online project, I would welcome feedback on the following (by Friday 26 August if possible). Of especial significance to help us make a final decision will be the amount of time required to deliver each option.

[REDACTED] has provided me with some usage statistics for the site, which I requested to determine the essentiality of the service. Deirdre gave me some comparative statistics for other products which the Library pays for.

Argus Online

2007 – 13,645 (8.48 pages per visit)
2008 – 19,192 (7.69 pages per visit)
2009 – 17,988 (7.88 pages per visit)
2010 – 21,231 (9.11 pages per visit)
2011 (to date) - 7,922 (6.82 pages per visit)

As previously mentioned, I believe the 2010 increase is due to some publicity about the impending index updates [http://blogs.abc.net.au/victoria/2010/05/the-argus-makes-a-return-online.html]. However, these statistics indicate a reasonable level of interest, the availability of the digitised pages not withstanding.

<u>eResources access</u>	To date this year
Onsite and offsite	

Subtotal via catalogue	11,406
Subtotal via eResources	52,501
Subtotal via Trove	10
Total	63,917

Options for proceeding

Option 1

In line with our strategic objective of reducing the number of separate websites the Library supports, we could treat the indexes as a source of data (like APAIS data) for Trove. Responsibilities to do this are suggested as follows:

CIB and CAB

- With assistance from ██████████, who is familiar with the data, confirm that the next set of index data can be mapped and loaded into Trove. Instead of being displayed as a separate service, integrate the display into Trove (could this be done in a way similar to the labelling being proposed for the AWW?). Could it also form a test case for other indexes?
- On load, the index entries should link to the relevant pages in Trove. A better level of granularity is unlikely because the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.

Trove support team

- With relevant IT staff, contact the index creators and propose the new approach

Web Publishing Branch

- Subject to confirmation, eventually switch off the old website.

Option 2

If it is not possible to load the data into Trove, then the likely actions are:

Web Publishing Branch

- Change the current website to reflect the availability of a new decade
- Provide a link to Trove to indicate the availability of fulltext.

CAB and CIB

- Load the new indexes using existing software, with assistance from ██████████

Trove support team

- Advise the index creators that data will be loaded only for the remainder of the decade, not to 2003 as proposed in the original grant.

Option 1 is preferred because:

- (a) it reduces the number of siloed websites hosted by the NLA
- (b) it adds value to Trove
- (c) it satisfies the request to be a part of Trove
- (d) new data can be added using a routine which does not have to be dusted off every few years or so.

Thanks for your advice,

[REDACTED]

[REDACTED]

[REDACTED]

Resource Sharing and Innovation Division
National Library of Australia

Parkes Place
Canberra ACT 2601

e: [REDACTED]

TRIMMED
DATE

From: [REDACTED]
Sent: Wednesday, 24 August 2011 1:45 PM
To: [REDACTED]
Subject: RE: making a decision about the Argus Online project

R11/50714
P11

Hi,

I had a look at a bit more sample data, to make this more meaningful, and double checked what I had done before. It looks like the page numbers are more reliable than I thought. (I had previously made a mistake with a couple of multi-page articles.) Out of 35 article references in the index, there was only one with an incorrect page, and another one that was missing altogether.



Argus.docx

[REDACTED]

From: [REDACTED]
Sent: Friday, 19 August 2011 11:16 AM
To: [REDACTED]
Subject: RE: making a decision about the Argus Online project

Hi,

Here's some information to start. A guess at accuracy of page numbers, and how many keywords from the index are searchable in Trove.

<< File: Argus.docx >>

[REDACTED]

From: [REDACTED]
Sent: Thursday, 18 August 2011 3:10 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: FW: making a decision about the Argus Online project

Hi [REDACTED]

Here is another piece of work for you. Background emails to follow.

Could you investigate please ([REDACTED] worked on the Argus index in the past) what the options are, issues would be, and work involved in adding this content into Trove. (A standout issue I think is that the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.)

Is a timeframe of 26 August achievable for you to do this? If not, when would be?

From: [redacted]
Sent: Thursday, 18 August 2011 2:50 PM
To: [redacted]
Cc: [redacted]
Subject: making a decision about the Argus Online project

[redacted]

Following on from the recent Trove/ANDP team discussion about continuing the Argus Online project, I would welcome feedback on the following (by Friday 26 August if possible). Of especial significance to help us make a final decision will be the amount of time required to deliver each option.

[redacted] has provided me with some usage statistics for the site, which I requested to determine the essentiality of the service. [redacted] gave me some comparative statistics for other products which the Library pays for.

Argus Online

- 2007 – 13,645 (8.48 pages per visit)
- 2008 – 19,192 (7.69 pages per visit)
- 2009 – 17,988 (7.88 pages per visit)
- 2010 – 21,231 (9.11 pages per visit)
- 2011 (to date) - 7,922 (6.82 pages per visit)

As previously mentioned, I believe the 2010 increase is due to some publicity about the impending index updates [<http://blogs.abc.net.au/victoria/2010/05/the-argus-makes-a-return-online.html>]. However, these statistics indicate a reasonable level of interest, the availability of the digitised pages notwithstanding.

eResources access

	To date this year
Onsite and offsite	
Subtotal via catalogue	11,406
Subtotal via eResources	52,501
Subtotal via Trove	10
Total	63,917

Options for proceeding

Option 1

In line with our strategic objective of reducing the number of separate websites the Library supports, we could treat the indexes as a source of data (like APAIS data) for Trove. Responsibilities to do this are suggested as follows:

CIB and CAB

- With assistance from [REDACTED], who is familiar with the data, confirm that the next set of index data can be mapped and loaded into Trove. Instead of being displayed as a separate service, integrate the display into Trove (could this be done in a way similar to the labelling being proposed for the AWW?). Could it also form a test case for other indexes?
- On load, the index entries should link to the relevant pages in Trove. A better level of granularity is unlikely because the indexes are derived from a copy of the Victorian version of the paper, whereas the digitised pages were created from a copy of the NSW published version.

Trove support team

- With relevant IT staff, contact the index creators and propose the new approach

Web Publishing Branch

- Subject to confirmation, eventually switch off the old website.

Option 2

If it is not possible to load the data into Trove, then the likely actions are:

Web Publishing Branch

- Change the current website to reflect the availability of a new decade
- Provide a link to Trove to indicate the availability of fulltext.

CAB and CIB

- Load the new indexes using existing software, with assistance from [REDACTED]

Trove support team

- Advise the index creators that data will be loaded only for the remainder of the decade, not to 2003 as proposed in the original grant.

Option 1 is preferred because:

- (a) it reduces the number of siloed websites hosted by the NLA
- (b) it adds value to Trove
- (c) it satisfies the request to be a part of Trove
- (d) new data can be added using a routine which does not have to be dusted off every few years or so.

Thanks for your advice,

[REDACTED]

[REDACTED]

[REDACTED]
Resource Sharing and Innovation Division
National Library of Australia

Parkes Place
Canberra ACT 2601

e: [REDACTED]

Analysis of Argus Index Data

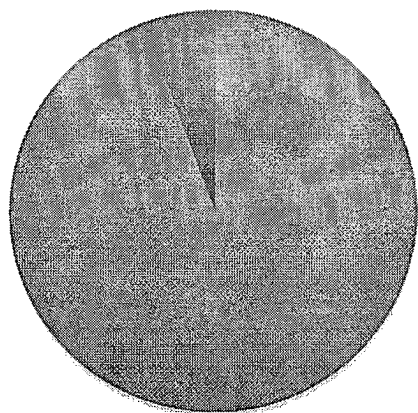
Overview

35 articles listed under four different headings in the Argus index were reviewed.

- 4 articles about “Aarons, Joseph” from 1870-1
- 11 articles about “Pacific Islands” from 1875
- 8 articles about “Waygunyah” from 1879
- 12 articles about “Weather” from 1873

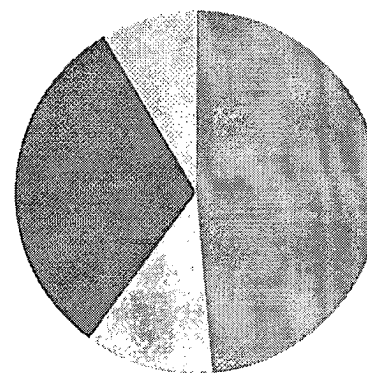
From this sample, the page number correlated between the index and the digitised article online for almost all index terms. About one third of online articles were missing a major keyword from the index. At least three of the four indexes were missing some relevant articles found by searching online.

Accuracy of page numbers in the Argus index



- Correct
- Incorrect

Articles containing keywords from index



- Contains all keywords
- Contains all major keywords
- Missing a major keyword
- Contains all keywords, but not all are searchable (yet)

Examples of missing minor keywords:

- Creation
- Capt (article contains “captain” instead)
- Complain (article contains complaint instead)

Aarons, Joseph

<http://www.nla.gov.au/apps/argus?action=BrowseSingleHeading&showOnly=AARONS,%20JOSEPH&year=1870&suppressPseud=n>

Argus Index				Trove				
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords
complains of nonreceipt of funds raised for masonic charities	1 Feb 1870	5	b	http://trove.nla.gov.au/ndp/del/page/225249 (article starts on pg 4)	5	b	N	complain (complaint) nonreceipt (receipt)
election for Latrobe Ward	10 Feb 1870	5	a	http://trove.nla.gov.au/ndp/del/page/225380	5	a	Y	
election for Latrobe Ward	11 Feb 1870	5	c	http://trove.nla.gov.au/ndp/del/page/225396 (article starts on pg 4)	5	c	Y*	
Haydamack lunch	3 Apr 1871	6	f[letter]	http://trove.nla.gov.au/ndp/del/page/229661	6	f	Y	

* An article better matching these search terms is found on pg 8, but is an election notice and likely to have been excluded from the Argus Index.

Some articles found online are missing from this index:

Election notices: <http://trove.nla.gov.au/ndp/del/article/5812153?searchTerm=Aarons%20joseph> "ELECTORB of LATROBE WARD,
Allow mo to oxproBs my sincere acknowledge- ment for the honour conferred In olcctlng mo your representativo in the City Council.
' Whilst I am permitted to onjoy thit privilege no exertion will bo wanting to further tho interests of your ward and the city.
Yours most obedlontly,
JOSEPH AARONS."

Pacific Islands

<http://www.nla.gov.au/apps/argus?action=BrowseIndex&year=1875&startFrom=p>

Argus Index				Trove				
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords
Capt. Macauley, Margaret and Jane, assaults female islander	2 Jan 1875	5	d	http://trove.nla.gov.au/ndp/del/page/239310	5	d	N	capt (captain)
King Kalakana, Sandwich Islands, visits Vic.	10 Feb 1875	10	d	http://trove.nla.gov.au/ndp/del/page/239908	10	d	Y	
Commodore J. G. Goodenough murdered by natives of Santa Cruz	24 Aug 1875	4	f [editorial]	http://trove.nla.gov.au/ndp/del/page/242956	4	g	N	pacific islands
“	24 Aug 1875	4	g	http://trove.nla.gov.au/ndp/del/page/242956	4	g	N	pacific islands
“	24 Aug 1875	5	g	http://trove.nla.gov.au/ndp/del/page/242957	5	g	N	Murdered J. G.
“	24 Aug 1875	6	a	http://trove.nla.gov.au/ndp/del/page/242959	6	a	N	J. G. (Jas. Graham) Pacific
“	28 Aug 1875	7	e	http://trove.nla.gov.au/ndp/del/page/243025	7	g	N	Murdered Natives Santa Cruz
“	30 Aug 1875	7	d	http://trove.nla.gov.au/ndp/del/page/243049	7	d	N	Murdered Natives Santa Cruz J. G. (James Graham)
“	8 Oct 1875	5	g	http://trove.nla.gov.au/ndp/del/page/241476	5	g	N	Murdered Natives Santa Cruz
“	15 Oct 1875	6	d	http://trove.nla.gov.au/ndp/del/page/241497 (article starts on pg 5)	6	d	Y	Cruz (Cru/,)
“	16 Oct 1875	4	c	http://trove.nla.gov.au/ndp/del/page/241604	4	c	Y	

Some articles found online are missing from this index:

31 August 1875 7a <http://trove.nla.gov.au/ndp/del/article/11522960> *"The death of Commodore Goodenough of which you have already received the full particulars has caused unfeigned sorrow in Sydney..."*

6 September 1875 1Sa <http://trove.nla.gov.au/ndp/del/article/7419003> *"H.M.S. Pearl arrived in Port Jackson, with the dead bodies of Commodore Goodenough and two seamen belonging to his vessel on board, all three having fallen victims to the treachery of the Bavares of Santa Cruz..."*

9 Oct 1875 6b <http://trove.nla.gov.au/ndp/del/article/7421860> *"The news of the murder of Commodore Goodenough caused great sorrow and mourning at Levuka. Seldom if ever is the loss of a man so universally felt..."*

Wahgunyah

<http://www.nla.gov.au/apps/argus?action=BrowseIndex&year=1879&startFrom=WAGES&headingType=%20&inc=n>

Argus Index				Trove					
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords	
opening of line from Springs to Wahgunyah	16 Jan 1879	5	a	http://trove.nla.gov.au/ndp/del/page/251508	5	a	Y	Wahgunyah (Wah"vmjah)	
opening of line from Springs to Wahgunyah	27 Jan 1879	5	a	http://trove.nla.gov.au/ndp/del/page/251674	5	a	Y	Wahgunyah (Waygunsah)	
opening of line from Springs to Wahgunyah	30 Jan 1879	5	e	http://trove.nla.gov.au/ndp/del/page/251721	5	e	Y		
creation as licensing district	3 Sep 1879	5	b	http://trove.nla.gov.au/ndp/del/page/252911	5	b	Y		
creation as licensing district	5 Sep 1879	7	c [letter]	http://trove.nla.gov.au/ndp/del/page/252959	7	c	N	creation	
creation as licensing district	9 Sep 1879	7	f [letter]	http://trove.nla.gov.au/ndp/del/page/253010	7	f	N	creation	
pleuro-pneumonia	18 Nov 1879	7	b	http://trove.nla.gov.au/ndp/del/page/254099	7	b	Y		
pleuro-pneumonia	1 Dec 1879	5	b	http://trove.nla.gov.au/ndp/del/page/254287	6	g	Y		

Some articles found online are missing from this index:

6 Feb 1879 7a Injury to Wahgunyah railway <http://trove.nla.gov.au/ndp/del/article/5931330?searchTerm=wahgunyah>

31 Oct 1879 3c Flood-waters at Wahgunyah <http://trove.nla.gov.au/ndp/del/article/5961491?searchTerm=wahgunyah>

29 Oct 1879 Purchase of Wahgunyah bridge <http://trove.nla.gov.au/ndp/del/article/5961166?searchTerm=wahgunyah>

Weather

<http://www.nla.gov.au/apps/argus?action=BrowseIndex&manualSearch=y&year=1873&startFrom=weather&go.x=22&go.y=18>

Argus Index				Trove				
Text	Date	Page	Col	Link to page indicated	Page	Col	Keywords in article?	Missing keywords
William Tennent Dawson, Gippsland surveyor, dies from intense heat on ride from Port Albert to Sale	22 Jan 1873	5	g	http://trove.nla.gov.au/ndp/del/page/234559	5	g	N	William Tennent (W.T). Gippsland Intense Heat ride
William Tennent Dawson, Gippsland surveyor, dies from intense heat on ride from Port Albert to Sale	25 Jan 1873 supp	1	g	http://trove.nla.gov.au/ndp/del/page/234610	?	?	N	All
chimney at Albion Hotel, Prahran, damaged by lightning strike	7 Feb 1873	5	b	http://trove.nla.gov.au/ndp/del/page/234812 (article starts on pg 4)	5	b	Y	
heavy rain and lightning storms in Melbourne	7 Feb 1873	5	b	http://trove.nla.gov.au/ndp/del/page/234812 (article starts on pg 4)	5	b	Y	
Victoria Racing Club wall struck down by lightning	7 Feb 1873	5	b	http://trove.nla.gov.au/ndp/del/page/234812 (article starts on pg 4)	5	b	Y	
thunderstorm at Brunswick and Coburg	10 Feb 1873	5	d	http://trove.nla.gov.au/ndp/del/page/234849 (article starts on pg 4)	5	d	Y	
violent storm at Castlemaine	11 Feb 1873	6	g	http://trove.nla.gov.au/ndp/del/page/234867	6	g	N*	Violent
severe hailstorm at Morang	13 Feb 1873	5	d	http://trove.nla.gov.au/ndp/del/page/234897 (article starts on pg 4)	5	d	Y	
thunderstorm at Warrnambool	19 Feb 1873	5	f	http://trove.nla.gov.au/ndp/del/page/234984	5	f	Y	
two escape lightning strike near Gibney's bridge	19 Feb 1873 supp	2	d	http://trove.nla.gov.au/ndp/del/page/234995	2	d	N	strike
violent thunderstorm in Melbourne	19 Feb 1873	5	f	http://trove.nla.gov.au/ndp/del/page/234984	5	f	Y	

severe storm damage at Inglewood	20 Feb 1873	5	e	http://trove.nla.gov.au/ndp/del/page/235005 (article starts on pg 4)	5	e	Y	
-------------------------------------	-------------	---	---	---	---	---	---	--

* An article better matching these search terms was found on pg 4, but it was not the correct article.

Note two different indexes are included for the same event:

- William Tennent Dawson, Gippsland district surveyor, dies from intense heat on long ride
(This one includes an extra article when compared to the following index. The article could not be found, so perhaps was a mistake.)
- William Tennent Dawson, Gippsland surveyor, dies from intense heat on ride from Port Albert to Sale

[REDACTED]
From: [REDACTED]
Sent: Friday, 26 August 2011 8:28 AM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

No: 211/80725

[REDACTED]
A solution, estimates, and delivered by the date I was hoping for. I am very grateful for that. You have given me generous detail to share with the project group as well.

[REDACTED] sent out advice this week indicating that many of our siloed services, including this one, need to be decommissioned by the end of June 2012. While the Argus Online can't take priority over Picture Australia, Music Australia and the API, that's the indicative timeframe for it.

Many thanks,
[REDACTED]

From: [REDACTED]
Sent: Thursday, 25 August 2011 7:13 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

Hi [REDACTED]

I've updated the document with some of your suggestions. [REDACTED]

For Option 1, Hiding non-matching terms sounds like customised behaviour that is a departure from any functionality that Trove has now. I agree it would be nice to highlight the matching term.

For Option 3, I thought that the tags would only be broken up when added manually in future. The existing index terms would be added without alteration. If G & J tried to edit them, they might not be able to save their changes – they'd have to break them up. It might not be a good idea anyway because articles tagged with one long term wouldn't be related to articles tagged with several smaller terms.

I've also added rough estimates, which [REDACTED] provided, to the summary table.

1. Add a record to Trove for each major heading (and display within Trove) **(3 weeks)**
2. Add index annotations to the digitised article in Trove **(2 weeks (or 1 week if not editable))**
3. Add index data to the digitised article in Trove, as tags **(1 week)**
4. In addition to solution 1, add a specialised interface for viewing the index "powered by Trove" **(2 weeks plus time for solution 1)**
5. Retain the existing stand alone app and load the new data there. **(1 week)**

From: [REDACTED]
Sent: Thursday, 25 August 2011 4:21 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

Thanks [REDACTED] maybe I'm getting too low level in these comments:

Option 1:

- The user searched for "sea porcupine" and then clicked on the Argus Index Warnambool link. I think by default it should hide all subheadings except those matching the search term, but make all subheadings easily viewable (rather than just showing all sub headings by default). This would help see all matching references across each year in the index. (If the user didn't have a search term, show everything!)
- Because we are linking to articles, I think we should also show the article name in the link, and maybe the article type (redundant?) and the length eg

16 March 1871, p 5 "HMAS Misery sunk in collision with marine invertebrate" [Article, 155 words]

- It should be possible to link to 2nd and subsequent pages of a multipage article as Newspapers URL exposes this, eg, 2nd page of 2 page article

<http://trove/ndp/del/article/7419558/241064?searchTerm=john%20olsen#pstart241064>

But these articles are very long, and even being on the right page still leaves a lot of work to do to find the reference. Argus does provide the column, and in theory we could use these to better guess and position the display to a column (newspapers currently does not support this in a URL, but could..), or, ideally, we could analyse the text of the article and the text of in the column and guess the "best" place to position the user.

Option 2:

- I like the navigation options! I think it would be nice if clicking on the link took you to the Argus index display in Option 1, showing the subheading you'd clicked on over the years and hiding the other sub headings (but making them available to view)
- The big problem with this is that as a naive user, I'd think "index terms" wasn't restricted to a decade or two of the Argus, and I'd be confused about it being so mysteriously sparse and selective and hence not representative of what is held in newspapers on that topic – actually, I'd hate it for that! Maybe if the heading was "Index Terms for The Argus 1870 – 1879" I'd be more understanding or something..

Option 3:

- You propose breaking up a tag such as "*Melbourne Public Library displays Denkmuler aus Aegypten und Aethiopien by Richard Lepsius, Monumens Egyptiens by C. Leemans, and Holy Land, Syria, Idumea, Arabia, Egypt and Nubia by David Robert*" into logical pieces, but how is this done? By Geraldine or by a program or by volunteers?
- Maybe Option 3 (which really shatters the integrity of the argus online index) is fine as long as Option 1 is also implemented (which keeps it for anyone who cares)?

In a separate email to [REDACTED], I've speculated that when a user is searching newspapers, then at the top of the related searches right hand column, we could show Argus Index results, helping to put them "in the flow" of the average newspaper user, who won't be looking for Argus Index results in some separate zone. I think that approach could co-exist with these options.

[REDACTED]

From: [REDACTED]
Sent: Wednesday, 24 August 2011 8:44 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: Mockups for each Argus index Option

Hi,

I have written some more detail and created mockups for each of the options for getting the Argus Index into Trove [REDACTED]

I also had a look for the longest index term (to see how it would go as a tag). One of the longest terms found in the index is 292 characters. With markup removed, it is 195 characters long.

Melbourne Public Library displays <properName>Denkmuler aus Aegypten und Aethiopien</properName> by Richard Lepsius, <properName>Monumens Egyptiens</properName> by C. Leemans, and <properName>Holy Land, Syria, Idumea, Arabia, Egypt and Nubia</properName> by David Robert

It looks like there are about 174,000 index terms (way to many to appear in a single tag cloud, for instance) I found this by running the following query, so Kent let me know if I didn't do it right: `select count (distinct displayname) from argheading`

[REDACTED]

[REDACTED]

[REDACTED] | IT Division | National Library of Australia

Jan Harris

TRIMmed

DATE:

No: R11/80725

From: [REDACTED]
Sent: Friday, 26 August 2011 4:29 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

04

Hi,

This extra feature adds about 0.5 weeks to option 1.

From: [REDACTED]
Sent: Friday, 26 August 2011 12:05 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

Thanks [REDACTED] – I just want to record a suggestion for putting Argus Index hits “in the flow” if the newspaper user, but showing them up the top right of the newspapers search results page (and that page only) – I’m presuming that if newspaper users search for a term that matches an Argus Index result, they’ll be interested in that. I know it looks like a crass add, but it’s the thought that counts..

<< File: olsen.jpg >>

From: [REDACTED]
Sent: Thursday, 25 August 2011 7:13 PM
To: [REDACTED]; [REDACTED]; [REDACTED]; Debbie Campbell
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

Hi [REDACTED]

I’ve updated the document with some of your suggestions. [REDACTED]

For Option 1, Hiding non-matching terms sounds like customised behaviour that is a departure from any functionality that Trove has now. I agree it would be nice to highlight the matching term.

For Option 3, I thought that the tags would only be broken up when added manually in future. The existing index terms would be added without alteration. If G & J tried to edit them, they might not be able to save their changes – they’d have to break them up. It might not be a good idea anyway because articles tagged with one long term wouldn’t be related to articles tagged with several smaller terms.

I’ve also added rough estimates, which [REDACTED] provided, to the summary table.

1. Add a record to Trove for each major heading (and display within Trove) (3 weeks)

2. Add index annotations to the digitised article in Trove (**2 weeks (or 1 week if not editable)**)
3. Add index data to the digitised article in Trove, as tags (**1 week**)
4. In addition to solution 1, add a specialised interface for viewing the index "powered by Trove" (**2 weeks plus time for solution 1**)
5. Retain the existing stand alone app and load the new data there. (**1 week**)

From: [REDACTED]
Sent: Thursday, 25 August 2011 4:21 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: RE: Mockups for each Argus index Option

Thanks [REDACTED], maybe I'm getting too low level in these comments:

Option 1:

- The user searched for "sea porcupine" and then clicked on the Argus Index Warnambool link. I think by default it should hide all subheadings except those matching the search term, but make all subheadings easily viewable (rather than just showing all sub headings by default). This would help see all matching references across each year in the index. (If the user didn't have a search term, show everything!)
- Because we are linking to articles, I think we should also show the article name in the link, and maybe the article type (redundant?) and the length eg

16 March 1871, p 5 "HMAS Misery sunk in collision with marine invertebrate" [Article, 155 words]

- It should be possible to link to 2nd and subsequent pages of a multipage article as Newspapers URL exposes this, eg, 2nd page of 2 page article

<http://trove/ndp/del/article/7419558/241064?searchTerm=john%20olsen#pstart241064>

But these articles are very long, and even being on the right page still leaves a lot of work to do to find the reference. Argus does provide the column, and in theory we could use these to better guess and position the display to a column (newspapers currently does not support this in a URL, but could..), or, ideally, we could analyse the text of the article and the text of in the column and guess the "best" place to position the user.

Option 2:

- I like the navigation options! I think it would be nice if clicking on the link took you to the Argus index display in Option 1, showing the subheading you'd clicked on over the years and hiding the other sub headings (but making them available to view)
- The big problem with this is that as a naive user, I'd think "index terms" wasn't restricted to a decade or two of the Argus, and I'd be confused about it being so mysteriously sparse and selective and hence not representative of what is held in newspapers on that topic – actually, I'd hate it for that! Maybe if the heading was "Index Terms for The Argus 1870 – 1879" I'd be more understanding or something..

Option 3:

- You propose breaking up a tag such as *"Melbourne Public Library displays Denkmuler aus Aegypten und Aethiopien by Richard Lepsius, Monumens Egyptiens by C. Leemans, and Holy Land, Syria, Idumea, Arabia, Egypt and Nubia by David Robert"* into logical pieces, but how is this done? By [REDACTED] or by a program or by volunteers?
- Maybe Option 3 (which really shatters the integrity of the Argus online index) is fine as long as Option 1 is also implemented (which keeps it for anyone who cares)?

In a separate email to Jo, I've speculated that when a user is searching newspapers, then at the top of the related searches right hand column, we could show Argus Index results, helping to put them "in the flow" of the average newspaper user, who won't be looking for Argus Index results in some separate zone. I think that approach could co-exist with these options.

[REDACTED]

From: [REDACTED]
Sent: Wednesday, 24 August 2011 8:44 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: Mockups for each Argus index Option

Hi,

I have written some more detail and created mockups for each of the options for getting the Argus Index into Trove. [REDACTED]

I also had a look for the longest index term (to see how it would go as a tag). One of the longest terms found in the index is 292 characters. With markup removed, it is 195 characters long.

Melbourne Public Library displays <properName>Denkmuler aus Aegypten und Aethiopien</properName> by Richard Lepsius, <properName>Monumens Egyptiens</properName> by C. Leemans, and <properName>Holy Land, Syria, Idumea, Arabia, Egypt and Nubia</properName> by David Robert

It looks like there are about 174,000 index terms (way to many to appear in a single tag cloud, for instance)

I found this by running the following query, so Kent let me know if I didn't do it right: `select count (distinct displayname) from argheading`

[REDACTED]

[REDACTED]

[REDACTED] | IT Division | National Library of Australia

your results:

- [us \(Melbourne, ... \(11,134\)](#)
- [dney Morning Herald...](#)
- [ast Australian \(Perth, ...](#)
- [urier-Mail \(Brisbane, ...](#)
- [g Bulletin \(Rockhampton, ...](#)
- [sbane Courier... \(2,524\)](#)
- [vertiser \(Adelaide, ... \(1,845\)](#)
- [vertiser \(Adelaide, ... \(1,563\)](#)
- [rcury \(Hobart, ... \(1,362\)](#)
- [gister \(Adelaide, ... \(1,235\)](#)
- [...](#)

[ry](#)
[sing \(23,477\)](#)
[\(13,094\)](#)
[d lists, results, guides](#)

[Notices \(3,456\)](#)
[ire \(4\)](#)
[ted](#)

[Count](#)

Digitised newspapers and more

Showing: 1 - 20 of **47,836** Refine search Sort by: Relevance

COURT JOHN OLSEN, I.O.F.
The Brisbane Courier (Qld. : 1864 - 1933) Monday 16 December 1929 p 19 Article
 ... COURT JOHN OLSEN, I.O.F. I Court John Olsen, Juvenile branch, of the Ancient Order of Foresters, held its annual breaking-up social on Friday night, in the Foresters' Hall, South Brisbane, there being a large attendance of children and parente. Bro..W. Reid (superintendent of the court) was ... 120 words

JOHN OLSEN MEMORIAL.
The Brisbane Courier (Qld. : 1864 - 1933) Tuesday 2 November 1926 p 3 Article
 ... JOHN OLSEN MEMORIAL. . The unveiling by Bro. J. Soden of the memorial stone erected in memory of the late **John Olsen** took place on Sunday afternoon last, in the South Brisbane Cemetery'. . Among- ... life lived by our departed friend, **John** .Olspu. **Mi. Olsen** lost his sight while dill a young man. Hl6 ... 310 words

THE LATE MR. JOHN OLSEN.
The Brisbane Courier (Qld. : 1864 - 1933) Friday 29 October 1926 p 3 Article
 ... THE LATE MR. **JOHN OLSEN**. Early in the current year committee of representatives ' of courts of ... Society for the Blind was formed with the view of erecting a me- morial to the late Mr. **John Olsen** (Junc- tion-terrace, Annerley). Mr. **Olsen**, who was blind, was endowed wth a character and personality out ... 183 words

MEMORIAL TABLET. LATE MR. JOHN OLSEN.
The Brisbane Courier (Qld. : 1864 - 1933) Thursday 4 November 1926 p 20 Article
 ... MEMORIAL TABLET. 'LATE MR. **JOHN OLSEN**. To perpetuate the memory of the late Mr. **John Olsen**, who ... black anti gilt lettering -"In affeetiollalo Inemori of **John Olsen**, founder of the QueeiiSlnnl Musical, ... said that the Inte Air **Olsen** had desotel his lue to the betteilient of blind peVc A.s n soung man ... 345 words

JOHN OLSEN'S DEATH. THE INQUEST. Perth, March 12.

Argus Index Results

Olsen, John

- obituary 14 Apr 1871
- fallen ill 9 Mar 1871
- never felt better 7 Mar 1871


Olsen, John Jnr

- forging and uttering 18 May 1871
- promotion to Chief Constable 30 June 1871

[more](#)

Books

view all **1,464** results



John Olsen / Deborah Hart
 Hart, Deborah, 1959-
Australian Painters of the Twentieth Century
 [Book, Article : 1991-2002]

[At National Library](#)

Options for ingesting the Argus Index into Trove

Last Edited: 29 August 2011

Table of Contents

Summary and Estimates	2
Option 1. Add a record to Trove for each major heading (e.g. Warrnambool).....	3
Option 2. Add index annotations (not tags or comments) to the digitised article	7
Option 3. Add index data to the digitised article as tags	8
Option 4. Create a specialised interface for viewing the index, powered by Trove.....	11
Option 5. In addition to Option 1, add an extra “Argus index” zone, when viewing newspaper articles ...	12

Summary and Estimates

Feature	Benefits	Challenges
<p>1. Add a record to Trove for each major heading (e.g. Warrnambool)</p> <p>ESTIMATE: 3 weeks</p>	<ul style="list-style-type: none"> Doesn't add clutter to NDP view Gets the data into Trove and searchable Preserves the structure of the sub-headings and the curated list of articles on a topic 	<ul style="list-style-type: none"> Will newspaper users find it? Is there something we can link to from the Argus title page, or elsewhere? In isolation, doesn't allow people looking at an article see the index/related articles A particular heading can still be browsed, but what about a year? Which zone? How to link to the right article? Requires a custom display to be written
<p>2. Add index annotations (not tags or comments) to the digitised article</p> <p>ESTIMATE: 2 weeks (1 week if not editable)</p>	<ul style="list-style-type: none"> Immediately accessible to newspapers users Can link to other articles with the same index term (reminiscent of tags), and to the (chronologically) next article Preserves the structure of the headings and curated list of articles. 	<ul style="list-style-type: none"> In isolation, might not preserve the structure of the index – no overview or ability to browse? How to avoid clutter/confusion on the article screen? Requires a custom display to be written How to identify the right article?
<p>3. Add index data to the digitised article as tags</p> <p>ESTIMATE: 1 week</p>	<ul style="list-style-type: none"> Immediately accessible to newspaper users Doesn't add new clutter to article screen – uses existing page sections Automatically provides links to other articles with the same tag/index term Provides an ongoing way to add more data 	<ul style="list-style-type: none"> How to identify the right article? Can't preserve sub-heading structure In isolation, might not preserve the structure of the index – no overview or ability to browse? (Could browse somewhat through tag cloud, especially if all tags start with a prefix.) Too long for existing tag structure, or can we work around this? (limit 60) Tags allow limited punctuation
<p>4. Create a specialised interface for viewing the index, powered by Trove (In addition to solution 1)</p> <p>ESTIMATE: 2 weeks plus time for solution 1</p>	<ul style="list-style-type: none"> Can browse by headings and by date Preserves the structure of the sub-headings and the curated list of articles on a topic More control over look and feel Retain separate brand 	<ul style="list-style-type: none"> Requires a custom display to be written The most work initially and ongoing Would records appear in the search results as well?
<p>5. Add an extra "Argus Index" zone (In addition to solution 1)</p> <p>ESTIMATE: 0.5 weeks Plus time for solution 1</p>	<ul style="list-style-type: none"> Preserves the structure of the index Puts the results in the flow of newspapers users 	<ul style="list-style-type: none"> Yet another zone – makes the results screen more complicated How to link to the right article? Index terms can't be seen when viewing the article.
<p>6. Retain the existing stand-alone application</p> <p>ESTIMATE: 1 week</p>	<ul style="list-style-type: none"> Preserves the structure of the index 	<ul style="list-style-type: none"> Trove users won't be able to see index terms and relationships A separate application must continue to be maintained

A combination of options might also be possible.

Work required to import data on an ongoing basis has not been considered.

Option 1. Add a record to Trove for each major heading (e.g. Warrnambool)

Users would search Trove and find results from all zones. One of these zones (yet to be determined) would include results from the Argus Index. (See fig. 1)

Trove
one search ... a wealth of information

All Books Pictures, photos, objects Journals, articles and data sets Digitised newspapers and more Music, sound and video Maps Diaries, letters, archives Archived websites (1996 – now) People and organisations Lists

sea porcupine Search

Available online Australian content In my libraries Advanced search Search tips

Refine your results:

Journals, articles and data sets Showing: 1 - 20 of 1,664 Sort by: Relevance

Format

- Article (1,664)
- Book chapter (6)
- Journal or magazine article (187)
- Other article (779)
- Report (0)
- Review (62)
- Database (0)
- Periodical (4)
- Journal, magazine, other (1)

Availability

- Online (1,664)
- Access conditions (1,209)
- Freely available (250)
- Unknown (1)

Decade

Language

Australian Content

Audience

Database

Books view all 2,073

The Deep-Sea Dredging Expedition in H.M.S. "Porcupine"
J. GWYN JEFFREYS
[Article : 1869]
Possibly online

Argus Index - Warrnambool
Geraldine Suter, John Hirst
[Article : 1870-1879]
Keywords: Warrnambool
... at port 5 May 1871 p.5 (deputation) didodon or sea porcupine found 18 May 1871 supp p.1 woollen manu ...
View online

Benthic dynamics at the carbonate mound regions of the Porcupine Sea Bight continental margin
White, Martin
International Journal of Earth Sciences
[Article : 2007]
Keywords: Earth sciences
... the carbonate mound provinces located at the Porcupine Bank/Sea Bight margin, NE Atlantic. The depth range ...
View online (conditions apply)

Reproductive strategies of two deep-sea gastropod species from the Porcupine Seabight (Northeast Atlantic)
Olabarria, C.; Ramirez-Llodra, E.

Figure 1. Example of the Argus index record for "Warrnambool", included in a list of Trove results.

There is no way to browse the index by date, jumping to a particular word in the index. Users can only search by keyword, and limit to a particular year (see fig 2.)

The screenshot shows the Trove search interface. At the top, the Trove logo is on the left, and navigation links (Home, About, Site news, Statistics, Tags, Help, Forum, Contact us) are on the right. Below the logo is the tagline "one search... a wealth of information". A search bar contains the text "argus index" and a "Search" button. To the right of the search bar are filters for "Available online", "Australian content", and "In my libraries". Below the search bar is a "Refine your results:" section with several filters: "Format" (Periodical, Journal, magazine, other, Newspaper, Article, Report), "Availability" (Online, Freely available), "Decade" (1870-1879), and "Year" (1879, 1878, 1877, 1876, 1875, 1874, 1873, 1872, 1871). The search results are displayed under the heading "Journals, articles and data sets" and show four entries, all for the year 1876. Each entry includes the title, author(s), and keywords, along with a "View online" link.

Home About Site news Statistics Tags Help Forum Contact us

one search... a wealth of information

It's free and it only takes a minute [Sign up](#) [Login](#)

All Books Pictures, photos, objects Journals, articles and data sets Digitised newspapers and more Music, sound and video Maps Diaries, letters, archives Archived websites (1996 - now) People and organisations

Lists

Search limited to: clear all
Decade: 1870-1879
Year: 1876

argus index Search

Available online Australian content In my libraries [Advanced search](#) [Search tips](#)

Refine your results: Journals, articles and data sets

Showing: 1 - 4 of 4 Sort by: Relevance

▼ Format
Periodical (4)
Journal, magazine, other (0)
Newspaper (0)
Article (1)
Report (1)

▼ Availability
Online (2)
Freely available (2)

▼ Decade
1870-1879 (4)

▼ Year
1879 (4)
1878 (4)
1877 (4)
1876 (4) **1876**
1875 (4)
1874 (4)
1873 (4)
1872 (4)
1871 (4)

The Argus Index - Weather
Geraldine Sutor, John Hirst
[Article: 1870-1879]
Keywords: Weather [View online](#)

The Argus Index - Zeplin, F.
Geraldine Sutor, John Hirst
[Article: 1870-1879]
Keywords: Zeplin, F. [View online](#)

The Argus Index - Abattoirs
Geraldine Sutor, John Hirst
[Article: 1870-1879]
Keywords: Abattoirs [View online](#)


The Argus Index - Oakleigh
Geraldine Sutor, John Hirst
[Article: 1870-1879]
Keywords: Oakleigh [View online](#)

Figure 2. Limiting a search to a particular year.

Clicking on this result (or the green “online” box) would display the full record for Warrnambool, within Trove (See fig. 3). On the full record screen:

- All sub-headings would be shown.
- The structure of sub-headings would be preserved.
- The article name could be shown, as well as the date and page number
- The sub headings could be either organised by year, or listed together, but not both. Users would not be able to choose which display is shown.
- If headings are organised by year:
 - As more years are added this may not scale well.
 - Users would have to scan each year for the term they are looking for.

NATIONAL LIBRARY OF AUSTRALIA



All	Books	Pictures, photos, objects	Journals, articles and data sets	Digitised newspapers and more	Music, sound and video	Maps	Diaries, letters, archives	Archived websites (1996 – now)	People and organisations	Lists
-----	-------	---------------------------	----------------------------------	-------------------------------	------------------------	------	----------------------------	--------------------------------	--------------------------	-------

The Argus Index - Warrnambool

Authors Geraldine Suter John Hirst

Bookmark <http://trove.nla.gov.au/work/3495788>

Work ID 3495788

Editions 1 version of this work is listed below

An index of articles published in The Argus newspaper between 1870 and 1879, for the term “Warrnambool”.

1870

meeting against selection on borough common
[10 Mar 1870 p.5 "THUESDAY MARCH 10, 1870." \[Article, 10413 words\]](#)

survey of harbour
[11 Mar 1870 p.5 "FRIDAY MARCH 11 1870" \[Article, 9701 words\]](#)
[18 Mar 1870 p.5, "FRIDAY MARCH 18 1870" \[Article, 9886 words\]](#)
[19 Dec 1870 p.5 "TUE IMPROVEMENT OF THE PORT ACCOMMODATION" \[Article, 1064 words\]](#)

banquet to Pharos officers
[15 Mar 1870 p.5 "TUESDAY MARCH 15 1870" \[Article, 9195 words\]](#)

search party for missing South Sea Islanders
[16 Mar 1870 p.7 "THE MEETING OF SHEEPOWNERS AT HAMILTON. TO THE EDITOR OF THE ARGUS. " \[Article, 694 words\]](#)

local land board deals with application for Woodford Forest allotments
[4 Jul 1870 p.5 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) WOOD'S POINT, SATURDAY." \[Article, 514 words\]](#)

laying of foundation stone at new masonic hall
[29 Aug 1870 p.6 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) QUEENSCLIFF, SATURDAY" \[Article, 1034 words\]](#)

floods
[13 Sep 1870 p.6 "THE FLOODS. GREAT DISTRESS AT ROCHESTER. \(FROM THE BENDIGO ADVERTISER, SEPT. 12.\) " \[Article, 2406 words\]](#)

excursion return trips to Melbourne
[29 Oct 1870 p.5 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) WAHGUNYAH, FRIDAY." \[Article, 1218 words\]](#)

Warrnambool directory and almanac for 1871
[28 Dec 1870 p.5 "WEDNESDAY, DECEMBER 28, 1870." \[Article, 7370 words\]](#)

Churches

Panmure Presbyterian church opened, designed by Andrew Kerr of Warrnambool
[7 Jan 1870 p.5 "FRIDAY, JANUARY 1, 1870" \[Article, 9238 words\]](#)

1871

potatoes successful
[24 Mar 1871 p.5 "FRIDAY, MARCH 24, 1871." \[Article, 9818 words\]](#)

funeral of Ernest Bostock
[28 Mar 1871 p.5 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) INGLEWOOD, MONDAY. ", \[Article, 622 words\]](#)

Not yet tagged

Add a tag

Not yet in any lists

[login to create your own lists](#)

Send to:

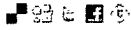


Figure 3. Viewing the Argus index entry for “Warrnambool” in Trove.

Links would lead to the article view (this appears feasible), rather than the page view. For multi-page articles, the correct page of the article would be shown, and possibly positioned on the right spot in the article. The terms in the index could be highlighted.

For the index term "laying of foundation stone at new Masonic hall 29 Aug 1870 p.6", the date and page would link to this article:

<http://trove.nla.gov.au/ndp/del/article/5830205?searchTerm=warrnambool+laying+foundation+stone+n+w+masonic+hall> (See fig. 4)

The screenshot shows the Trove website interface. At the top, there is a navigation bar with links for Home, About, Browse, Help, and a search bar containing the text 'warrnambool laying foundati'. Below this, the page title is 'The Argus (Melbourne, Vic. : 1848 - 1956) about' and the date is 'Monday 29 August 1870'. The page number is 'Page 6 of 8'. The main content area displays the newspaper text, which is organized into sections under the heading 'TELEGRAPHIC DESPATCHES'. The sections are: 'QUEENSLIFF, SATURDAY' (reporting on a shipwreck), 'BELFAST, SATURDAY' (reporting on a meeting), and 'WARRNAMBOOL, SATURDAY' (reporting on the laying of a foundation stone). The 'WARRNAMBOOL, SATURDAY' section is the most detailed and matches the search terms in the URL. On the left side of the page, there are various interactive elements like 'Lists (None yet)', 'Tagged (None yet)', and 'Comments (None yet)'. At the bottom right, there is a 'ZOOM' button.

Figure 4. After linking from the Argus index record in Trove to a particular article

Option 2. Add index annotations (not tags or comments) to the digitised article

Index terms would be added to the article, preserving the structure of the subheadings (See fig. 5). Clicking on a term could return other articles with the same index term. (But not articles tagged with the same term?) Users could navigate to the (chronologically) next article with the same index term.

Normal users couldn't add or edit index terms. Maybe [redacted] and [redacted], and perhaps some other users, could have special permissions to be able to add/edit/remove index terms?

Will users be confused as to what these are, why they are not editable, or why they appear only for such a selected subset of data?

Index terms would not be included in the tag cloud.

This solution depends on being able to automatically identify the correct article.

The screenshot shows the Trove website interface. At the top, the Trove logo is displayed with the tagline "Digitised newspapers and more". Navigation links for Home, About, Browse, and Help are visible. A search bar is present with a "Sign up" button and a "Login" button. The main content area displays a newspaper article from "The Argus (Melbourne, Vic. : 1848 - 1956)" dated Monday 29 August 1870. The article is titled "TELEGRAPHIC DESPATCHES." and contains two sections: "QUEENSCLIFF, SATURDAY." and "BELFAST, SATURDAY." The article text describes the towing of a wrecked barque and a meeting of the electors of the Western Provinces. On the left side of the article, there is a sidebar with various interactive elements: "Lists (None yet)", "Tagged (None yet)", "Index Terms" (with sub-headers "Warrnambool" and "Weather"), "Comments (None yet)", and "Electronically Translated Text". The "Index Terms" section shows terms like "laying of foundation stone at new masonic hall" and "heavy gale at Sydney". The article text is partially obscured by the sidebar and the image quality is somewhat grainy.

Figure 5. An article displaying index terms (in a similar way to tags).

Option 3. Add index data to the digitised article as tags

Index terms would be added to the article as tags (see fig. 6). Clicking on a term would return other articles with the same term or tag.

Other users couldn't edit or remove the terms, but could add new tags. [redacted] and [redacted] could remove or edit them using a particular (shared) user account.

The structure of sub-headings would not be preserved. Mark-up of proper names would be lost. Users could not navigate to the (chronologically) next article with the same index term except by clicking on the tag and sorting the results by date.

This solution relies on being able to identify the article the index refers to.

The screenshot shows a web browser displaying a digitized newspaper article. At the top, the 'Trove' logo is visible, along with navigation links for 'Home', 'About', 'Browse', and 'Help'. The article title is 'The Argus (Melbourne, Vic. : 1848 - 1956)' and the date is 'Monday 29 August 1870'. The article content includes the following text:

TELEGRAPHIC DESPATCHES.
(FROM OUR OWN CORRESPONDENTS.)
QUEENSCLIFF, SATURDAY.
The Pharos has towed the stern port quarter, and 50 feet of the deck, of the wrecked barque *Sussex* into Swan Bay.

BELFAST, SATURDAY.
Mr. P. J. Martin addressed a well-attended meeting of the electors of this division of the Western Province. He stated he was in favour of secular education in thinly-populated districts where one common school would meet the requirements of the inhabitants, but was in favour of the denominational system in the centres of population, where there were sufficient children to attend them. He was in favour of immigration; a five years' tenure of office for the members of the Upper House; the pink line of railway and harbour improvements. He was opposed to payment of members. He would like to see commercial men in the Upper House, for, had such been the case, no dead-lock would have ever occurred over the Darling grant. The candidate was well received.

The schooner *Mary Stewart* left the wharf this morning, there being 8ft. 6in. on the bar, making the trip in and out in perfect safety.

On the left side of the page, there is a sidebar with several sections:

- Lists (None yet)**: Includes a 'Login to create lists' button.
- Tagged**: Lists tags such as 'Warrnambool', 'Weather', 'laying of foundation stone at new masonic hall', and 'heavy gale at sydney'. Includes an 'Add New Tag' button.
- Comments (None yet)**: Includes an 'Add New Comment' button.
- Electronically Translated Text**: Includes a 'Fix this Text' button and a section for corrections.

At the bottom of the sidebar, it says 'On page 6' and repeats the article title and content.

Figure 6. Article with index terms added as tags

Index terms would be included in the tag cloud, which would provide some kind of browse option. To see only Argus index terms, users would need to either:

- search for the word “Argus” (And all tags would need to have the prefix “Argus”) (See fig 7.) OR
- view the “ArgusIndex” user profile (See fig. 8)

The Trove tag cloud can display up to 3,000 tags at any one time. The tag cloud can also be searched. The Argus index contains 174,323 different index terms at the moment.

The screenshot shows the Trove website interface. At the top, there is a navigation menu with links for Home, About, Site news, Statistics, Tags, Help, Forum, and Contact us. Below this is a search bar with a 'Search' button. The search results are displayed as a tag cloud, where the word 'ARGUS' is the most prominent. Other tags include 'ARGUS - Abbotair', 'ARGUS - Ballarat', 'ARGUS - Banalla', 'ARGUS - Birds', 'ARGUS - Cabs', 'ARGUS - Carlton', 'ARGUS - Cattle', 'ARGUS - Charities', 'ARGUS - Child abuse', 'ARGUS - Chinese', 'ARGUS - Church of England', 'ARGUS - Death', 'ARGUS - Diners and dining', 'ARGUS - Diptheria', 'ARGUS - Domain gardens', 'ARGUS - Drowning', 'ARGUS - Education', 'ARGUS - Education act', 'ARGUS - Education Department', 'ARGUS - Elections', 'ARGUS - Emerald Hill', 'ARGUS - Entertainers', 'ARGUS - Factory work', 'ARGUS - Finance', 'ARGUS - Fire control', 'ARGUS - Fires', 'ARGUS - Fisheries', 'ARGUS - Fishing', 'ARGUS - Fitzroy', 'ARGUS - France', 'ARGUS - Gambling', 'ARGUS - Geelong', 'ARGUS - Gold', 'ARGUS - Government House', 'ARGUS - Great Britain', 'ARGUS - Harrison, James', 'ARGUS - Health, Central board of', 'ARGUS - Higinbotham, George', 'ARGUS - Holidays', 'ARGUS - Homeless', 'ARGUS - Homicide', 'ARGUS - Immigration', 'ARGUS - Industrial accidents', 'ARGUS - Infanticide', 'ARGUS - Intercolonial relations', 'ARGUS - Judiciary', 'ARGUS - Kew', 'ARGUS - Labour', 'ARGUS - Land', 'ARGUS - Land acts', 'ARGUS - Land and Works, department of', 'ARGUS - Land selection', 'ARGUS - Larrikinism', 'ARGUS - Law', 'ARGUS - Legal profession', 'ARGUS - Licensing', 'ARGUS - Litigation', 'ARGUS - Local government', 'ARGUS - Maintenance (family)', 'ARGUS - Marriage', 'ARGUS - Meat preservation', 'ARGUS - Medical profession', 'ARGUS - Melbourne', 'ARGUS - National Agricultural Society of Victoria', 'ARGUS - New South Wales', 'ARGUS - Pacific Islands', 'ARGUS - Parks and gardens', 'ARGUS - Parliaments', 'ARGUS - Pastoral Industry', 'ARGUS - Patents and Inventions', 'ARGUS - Railway accidents', 'ARGUS - Railways', 'ARGUS - St Kilda', 'ARGUS - Sandhurst', 'ARGUS - Schools', 'ARGUS - Singers', 'ARGUS - Tariffs', 'ARGUS - Teaching', 'ARGUS - Telegraph', 'ARGUS - United States', 'ARGUS - Vandalism', 'ARGUS - Wages', 'ARGUS - Weather', 'ARGUS - Wetlands', and 'ARGUS - Yarra River'. At the bottom of the page, there is a footer with the National Library of Australia logo and a disclaimer.

Figure 7. Index terms in tag cloud, with “ARGUS” prefix

one search a wealth of information

All Books Pictures, photos, objects Journals, articles and data sets Digitised newspapers and more Music, sound and video Maps Diaries, letters, archives Archived websites (1996 – now) People and organisations Lists

Search

Available online Australian content In my libraries [Advanced search](#) [Search tips](#)

Information about Trove user: ArgusIndex

[View user profile in the Trove forum](#)

Tags: [Recent comments](#) [Text corrections](#) [Recent merge/splits](#) [Lists](#)

Display options

Containing: Display only: top tags Added: By: [Refresh](#)

Abolition Abusive Language **Accidents** Acclimatisation society **Actors** Alcohol abuse **Ballarat** Banalla Birds Cabs **Carlton** Cattle [Charities](#) Child abuse **Chinese** Church of England Death Diners and dining Diphtheria Domain gardens Drowning Education Education act **Education** Department Elections Emerald Hill **Entertainers** Factory work Finance Fire control **Fires** Fisheries **Fishing** Fitzroy France Gambling **Geelong** Gold Government House Great Britain Harrison, James **Health, Central board of** Higinbotham, George Holidays homeless **Homicide** Immigration Industrial accidents Intaricide Intercolonial relations Judiciary Kew **Labour** Land Land acts Land and Works, department of Land selection Laminium **Law** Legal profession Licensing **Litigation** Local government Maintenance (family) **Marriage** Meat preservation Medical profession **Melbourne** National Agricultural Society of Victoria **New South Wales** Pacific Islands Parks and gardens **Parliaments** Pastoral Industry **Patents and Inventions** Railway accidents **Railways** St Kilda Sandhurst schools Singers Tariffs **Teaching** Telegraph United States vandalism **Wages** Weather wetlands Yarra River

[Bookmark this tag cloud](#)



Figure 8. Index terms in tag cloud for user “ArgusIndex”

Trove tags do not allow punctuation other than apostrophe, hyphen, parentheses, comma and forward slash.

There is a 60 character limit on tags in Trove. In order to add new terms or edit existing ones, either:

- the limit would need to be increased for all users
- the limit would need to be increased for the ArgusIndex user
- indexing practice would need to be adjusted, by breaking index phrases into multiple tags. This may also achieve a greater hit rate of related articles.

For example, in future if trying to add the index term “Melbourne Public Library displays Denkmuler aus Aegypten und Aethiopian by Richard Lepsius, Monumens Egyptiens by C. Leemans, and Holy Land, Syria, Idumea, Arabia, Egypt and Nubia by David Robert” as a tag, it could be broken up in the following way:

- Melbourne public library
- Denkmuler aus Aegypten und Aethiopian by Richard Lepsius
- Monumens Egyptiens by C Leemans
- Holy Land, Syria, Idumea, Arabia, Egypt, Nubia by D Robert

An article with these tags would not be related to other articles tagged with the full phrase.

Option 4. Create a specialised interface for viewing the index, powered by Trove

The Argus index would still be available as a separate service, and is likely to retain similar functionality to the current service. This is the most labour intensive solution. (See fig. 9)

NATIONAL LIBRARY OF AUSTRALIA

The Argus.

HOME • BROWSE HEADINGS • BROWSE INDEX • SEARCH • HELP • CONTACT

Powered by Trove

THE ARGUS INDEX
1870 TO 1879

Geraldine Suter, Chief Indexer
John Hirst, Editor-in-Chief

© 2005

The Argus, published in Melbourne from 1846-1957, was the quality newspaper of record in the nineteenth and first part of the twentieth century. This index to the Argus covers the decade 1870-1879.

Think of this website as consisting of a core of ten separate and linked annual indexes. The indexes can be:

- browsed directly with the [Browse index](#) function
- accessed through a cumulated list of headings for all years with [Browse headings](#)
- approached through the [Search](#) function

The site does not provide online access to the full text of the paper. A separate page describes how to [get copies of articles](#) found through the index.

HOME • BROWSE HEADINGS • BROWSE INDEX • SEARCH • HELP • CONTACT

THE ARGUS INDEX

STORIES

- Fish - [English roach caught in Yarra River](#)
- Infanticide - [Daughter of Ellen Tighe found in cesspit](#)
- Kelly Gang - [the whole story as it happened](#)
- Marcus Clarke - [his natural life serialised](#)
- Art - [Eugene Von Guerard - Recent work](#)
- Science - [government explores scientists but often neglects to consult them](#)

LINKS

- [The Argus Index Project](#)
- [Getting copies of articles](#)
- [Indexes to the Argus for other years](#)

Figure 9. A specialised interface for the Argus Index, powered by Trove

Option 5. In addition to Option 1, add an extra "Argus index" zone, when viewing newspaper articles

When a user searches the newspapers zone in Trove, show them any hits on the Argus Index at the top right of the search results page (fig. 10). The zone wouldn't appear when looking at any other results screens.

Don't show the extra zone on any of the other results pages.

The screenshot shows the Trove search interface with the search term 'john olsen'. The search results are filtered to 'Digitised newspapers and more', showing 47,836 results. The main results list includes several newspaper articles from the Brisbane Courier and Western Argus, such as 'COURT JOHN OLSEN, I.O.F.', 'JOHN OLSEN MEMORIAL.', 'THE LATE MR. JOHN OLSEN.', 'MEMORIAL TABLET. LATE MR. JOHN OLSEN.', and 'JOHN OLSEN'S DEATH. THE INQUEST. Perth, March 12.'.

On the right side of the page, there is a sidebar titled 'Argus Index Results' which lists specific dates and events related to John Olsen:

- Olsen, John
 - obituary **14 Apr 1871**
 - fallen ill **9 Mar 1871**
 - never felt better **7 Mar 1871**
- Olsen, John Jnr
 - forging and uttering **18 May 1871**
 - promotion to Chief Constable **30 June 1871**

 A 'more' link is provided below the Argus Index Results.

Below the Argus Index Results, there is a 'Books' section showing search results for books related to John Olsen, including 'John Olsen / Deborah Hart' and 'John Olsen / by Virginia Spate'.

Figure 10. An extra "Argus Index" zone, only shown on the "Newspapers" results screen.

TRIMmed

From: [REDACTED]
Sent: Tuesday, 13 December 2011 2:07 PM
To: [REDACTED]
Cc: [REDACTED]
Subject: Argus Index Options
Attachments: ArgusIndex_Trove_Options.docx
Importance: High

DATE:

No: R11/113018

P13

Hi All,

Attached is a more up to date version of the Argus Index Options paper for our meeting this afternoon.

[REDACTED]

TRIMMED
DATE:
No: 12/11/2018

Options for ingesting the Argus Index into Trove

Last Edited: 29 August 2011

Table of Contents

Summary and Estimates.....	2
Option 1. Add a record to Trove for each major heading (e.g. Warrnambool)	3
Option 2. Add index annotations (not tags or comments) to the digitised article	7
Option 3. Add index data to the digitised article as tags	8
Option 4. Create a specialised interface for viewing the index, powered by Trove.....	11
Option 5. In addition to Option 1, add an extra “Argus index” zone, when viewing newspaper articles ...	12

Summary and Estimates

Feature	Benefits	Challenges
<p>1. Add a record to Trove for each major heading (e.g. Warrnambool)</p> <p>ESTIMATE: 3 weeks</p>	<ul style="list-style-type: none"> • Doesn't add clutter to NDP view • Gets the data into Trove and searchable • Preserves the structure of the sub-headings and the curated list of articles on a topic 	<ul style="list-style-type: none"> • Will newspaper users find it? • Is there something we can link to from the Argus title page, or elsewhere? • In isolation, doesn't allow people looking at an article see the index/related articles • A particular heading can still be browsed, but what about a year? • Which zone? • How to link to the right article? • Requires a custom display to be written
<p>2. Add index annotations (not tags or comments) to the digitised article</p> <p>ESTIMATE: 2 weeks (1 week if not editable)</p>	<ul style="list-style-type: none"> • Immediately accessible to newspapers users • Can link to other articles with the same index term (reminiscent of tags), and to the (chronologically) next article • Preserves the structure of the headings and curated list of articles. 	<ul style="list-style-type: none"> • In isolation, might not preserve the structure of the index – no overview or ability to browse? • How to avoid clutter/confusion on the article screen? • Requires a custom display to be written • How to identify the right article?
<p>3. Add index data to the digitised article as tags</p> <p>ESTIMATE: 1 week</p>	<ul style="list-style-type: none"> • Immediately accessible to newspaper users • Doesn't add new clutter to article screen – uses existing page sections • Automatically provides links to other articles with the same tag/index term • Provides an ongoing way to add more data 	<ul style="list-style-type: none"> • How to identify the right article? • Can't preserve sub-heading structure • In isolation, might not preserve the structure of the index – no overview or ability to browse? (Could browse somewhat through tag cloud, especially if all tags start with a prefix.) • Too long for existing tag structure, or can we work around this? (limit 60) • Tags allow limited punctuation
<p>4. Create a specialised interface for viewing the index, powered by Trove (In addition to solution 1)</p> <p>ESTIMATE: 2 weeks plus time for solution 1</p>	<ul style="list-style-type: none"> • Can browse by headings and by date • Preserves the structure of the sub-headings and the curated list of articles on a topic • More control over look and feel • Retain separate brand 	<ul style="list-style-type: none"> • Requires a custom display to be written • The most work initially and ongoing
<p>5. Retain the existing stand-alone application</p> <p>ESTIMATE: 1 week</p>	<ul style="list-style-type: none"> • Preserves the structure of the index 	<ul style="list-style-type: none"> • Trove users won't be able to see index terms and relationships • A separate application must continue to be maintained
<p>6. Add an extra "Argus Index" zone (In addition to solution 1)</p> <p>ESTIMATE: 0.5 weeks Plus time for solution 1</p>	<ul style="list-style-type: none"> • Preserves the structure of the index • Puts the results in the flow of newspapers users 	<ul style="list-style-type: none"> • Yet another zone – makes the results screen more complicated • How to link to the right article? • Index terms can't be seen when viewing the article.

A combination of options might also be possible.

Option 1. Add a record to Trove for each major heading (e.g. Warrnambool)

Users would search Trove and find results from all zones. One of these zones (yet to be determined) would include results from the Argus Index. (See fig. 1)

The screenshot shows the Trove search interface. At the top, there are navigation tabs for different content types: All, Books, Pictures, photos, objects, Journals, articles and data sets, Digitised newspapers and more, Music, sound and video, Maps, Diaries, letters, archives, Archived websites (1996 - now), People and organisations, and Lists. The search bar contains the text 'sea porcupine' and a 'Search' button. Below the search bar, there are checkboxes for 'Available online', 'Australian content', and 'In my libraries', along with links for 'Advanced search' and 'Search tips'.

The main content area is titled 'Journals, articles and data sets' and shows 'Showing: 1 - 20 of 1,664' results. The results are sorted by 'Relevance'. The first result is 'The Deep-Sea Drifing Expedition in H.M.S. "Porcupine"' by J. GWYN JEFFREYS, dated 1869. The second result is 'Argus Index - Warrnambool' by Geraldine Suter, John Hirst, dated 1870-1879. A large black arrow points to this entry. The third result is 'Benthic dynamics at the carbonate mound regions of the Porcupine Sea Bight continental margin' by White, Martin, dated 2007. The fourth result is 'Reproductive strategies of two deep-sea gastropod species from the Porcupine Sea Bight (Northeast Atlantic)' by Olabarria, C.; Ramirez-Llodra, E.

On the left side, there is a 'Refine your results:' section with filters for 'Format', 'Availability', 'Decade', 'Language', 'Australian Content', 'Audience', and 'Database'. The 'Format' filter is expanded, showing options like 'Article (1,595)', 'Book chapter (35)', 'Journal or magazine article (830)', 'Other article (575)', 'Report (13)', 'Review (82)', 'Periodical (4)', and 'Journal, magazine, other (1)'. A large black arrow points to the 'Journal, magazine, other (1)' option.

On the right side, there is a 'Books' section with a link to 'view all 2,973' items.

Figure 1. Example of the Argus index record for "Warrnambool", included in a list of Trove results.

There is no way to browse the index by date, jumping to a particular word in the index. Users can only search by keyword, and limit to a particular year (see fig 2.)


The screenshot shows the Trove website interface. At the top, there is a navigation bar with links for Home, About, Site news, Statistics, Tags, Help, Forum, and Contact us. Below this is a search bar with the text 'one search... a wealth of information' and buttons for 'Sign up' and 'Login'. A horizontal menu lists various content types: All, Books, Pictures, photos, objects, Journals, articles and data sets (highlighted), Digitised newspapers and more, Music, sound and video, Maps, Diaries, letters, archives, Archived websites (1996 - now), and People and organisations. Below the menu is a search bar containing 'argus index' and a 'Search' button. To the left of the search bar, there are filters for 'Decade: 1870-1879' and 'Year: 1876'. Below the search bar, there are checkboxes for 'Available online', 'Australian content', and 'In my libraries', along with links for 'Advanced search' and 'Search tips'. On the left side, there is a 'Refine your results:' section with expandable categories: Format (Periodical (4), Journal, magazine, other (2), Newspaper (2), Article (1), Report (1)), Availability (Online (2), Freely available (2)), Decade (1870-1879 (4)), and Year (1879 (4), 1878 (4), 1877 (4), 1876 (4), 1875 (4), 1874 (4), 1873 (4), 1872 (4), 1871 (4)). The main content area is titled 'Journals, articles and data sets' and shows 'Showing: 1 - 4 of 4' results. The results are sorted by 'Relevance'. Each result entry includes the title, authors, article information, keywords, and a 'View online' button. The results are: 1. 'The Argus Index - Weather' by Geraldine Sutor, John Hirst [Article: 1870-1879], Keywords: Weather. 2. 'The Argus Index - Zeplin, F.' by Geraldine Sutor, John Hirst [Article: 1870-1879], Keywords: Zeplin, F. 3. 'The Argus Index - Abattoirs' by Geraldine Sutor, John Hirst [Article: 1870-1879], Keywords: Abattoirs. 4. 'The Argus Index - Oakleigh' by Geraldine Sutor, John Hirst [Article: 1870-1879], Keywords: Oakleigh.

Figure 2. Limiting a search to a particular year.

Clicking on this result (or the green "online" box) would display the full record for Warrnambool, within Trove (See fig. 3). On the full record screen:

- All sub-headings would be shown.
- The structure of sub-headings would be preserved.
- The article name could be shown, as well as the date and page number
- The sub headings could be either organised by year, or listed together, but not both. Users would not be able to choose which display is shown.
- If headings are organised by year:
 - As more years are added this may not scale well.
 - Users would have to scan each year for the term they are looking for.

NATIONAL LIBRARY OF AUSTRALIA



one search ... a wealth of information

All	Books	Pictures, photos, objects	Journals, articles and data sets	Digitised newspapers and more	Music, sound and video	Maps	Diaries, letters, archives	Archived websites (1996 – now)	People and organisations	Lists
-----	-------	---------------------------	----------------------------------	-------------------------------	------------------------	------	----------------------------	--------------------------------	--------------------------	-------

The Argus Index - Warrnambool

Authors: [Geraldine Suter, John Hirst](#)

Bookmark: <http://trove.nla.gov.au/work/3495788>

Work ID: 3495788

Editions: 1 version of this work is listed below

An index of articles published in The Argus newspaper between 1870 and 1879, for the term "Warrnambool".

1870

meeting against selection on borough common
[10 Mar 1870 p.5 "THUESDAY, MARCH 10, 1870." \[Article, 10413 words\]](#)

survey of harbour
[11 Mar 1870 p.5 "FRIDAY MARCH 11 1870" \[Article, 9701 words\]](#)
[18 Mar 1870 p.9 "FRIDAY MARCH 18 1870" \[Article, 9886 words\]](#)
[19 Dec 1870 p.5 "TUE IMPROVEMENT OF THE PORT ACCOMMODATION" \[Article, 1064 words\]](#)

banquet to Pharos officers
[15 Mar 1870 p.5 "TUESDAY MARCH 15 1870" \[Article, 5195 words\]](#)

search party for missing South Sea Islanders
[16 Mar 1870 p.7 "THE MEETING OF SHEEPOWNERS AT HAMILTON, TO THE EDITOR OF THE ARGUS." \[Article, 694 words\]](#)

local land board deals with application for Woodford Forest allotments
[4 Jul 1870 p.5 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) WOOD'S POINT, SATURDAY." \[Article, 514 words\]](#)

laying of foundation stone at new masonic hall
[29 Aug 1870 p.6 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) QUEENSCLIFF, SATURDAY" \[Article, 1034 words\]](#)

floods
[19 Sep 1870 p.6 "THE FLOODS. GREAT DISTRESS AT ROCHESTER. \(FROM THE BENDIGO ADVERTISER, SEPT. 12.\)" \[Article, 2406 words\]](#)

excursion return trips to Melbourne
[29 Oct 1870 p.5 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) WAHGUNYAH, FRIDAY." \[Article, 1218 words\]](#)

Warrnambool directory and almanac for 1871
[28 Dec 1870 p.5 "WEDNESDAY, DECEMBER 28, 1870." \[Article, 7370 words\]](#)

Churches
[Panmure Presbyterian church opened, designed by Andrew Kerr of Warrnambool](#)
[7 Jan 1870 p.5 "FRIDAY, JANUARY 1, 1870" \[Article, 9235 words\]](#)

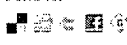
Not yet tagged

Add a tag

Not yet in any lists

login to create your own lists

Send to:



1871

potatoes successful
[24 Mar 1871 p.5 "FRIDAY, MARCH 24, 1871." \[Article, 9818 words\]](#)

funeral of Ernest Bostock
[28 Mar 1871 p.5 "TELEGRAPHIC DESPATCHES. \(FROM OUR OWN CORRESPONDENTS.\) INGLEWOOD, MONDAY." \[Article, 622 words\]](#)

Figure 3. Viewing the Argus index entry for "Warrnambool" in Trove.

Links would lead to the article view (this appears feasible), rather than the page view. For multi-page articles, the correct page of the article would be shown, and possibly positioned on the right spot in the article. The terms in the index could be highlighted.

For the index term "laying of foundation stone at new Masonic hall 29 Aug 1870 p.6", the date and page would link to this article:

<http://trove.nla.gov.au/ndp/del/article/5830205?searchTerm=warrnambool+laying+foundation+stone+new+masonic+hall> (See fig. 4)

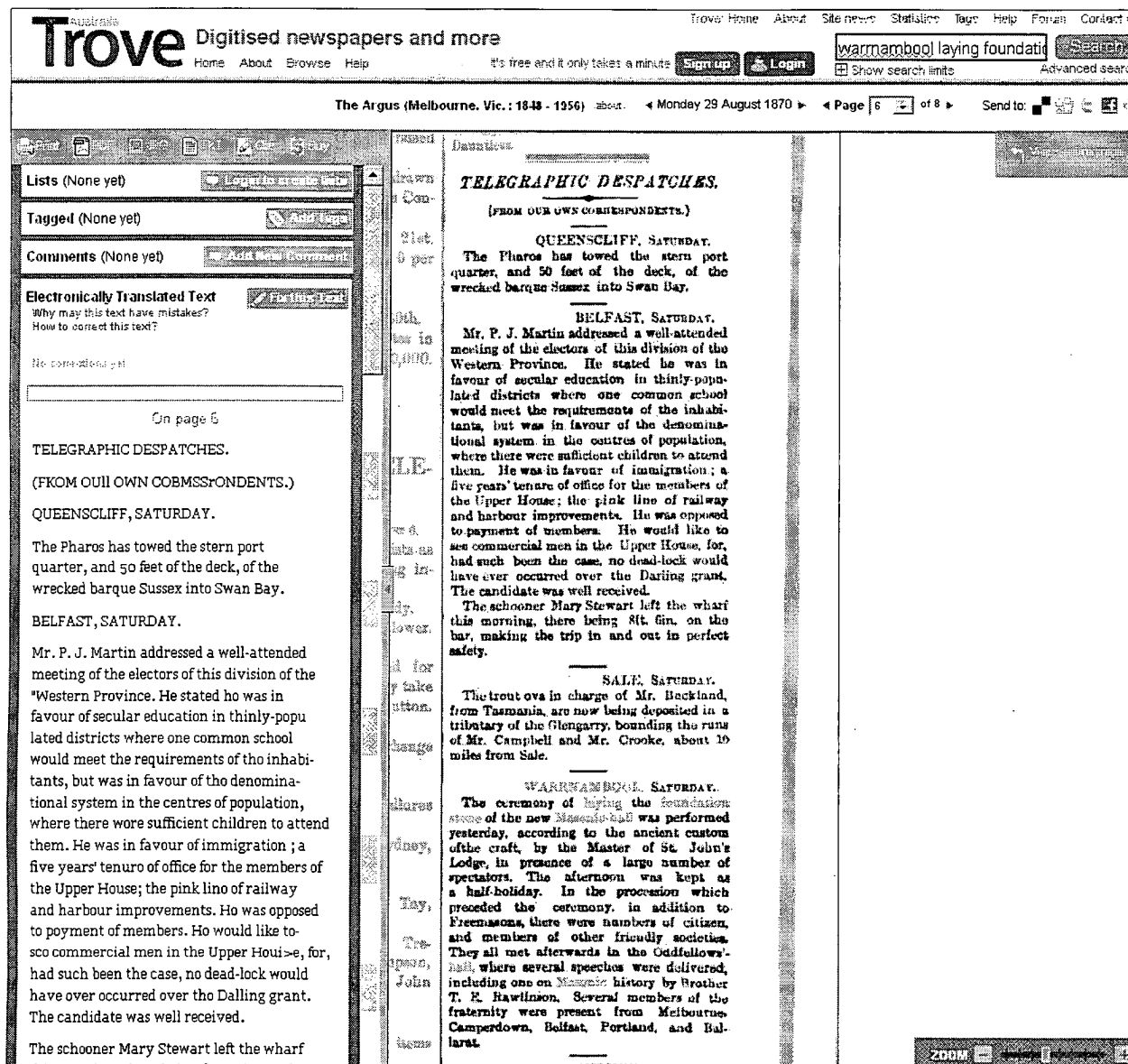


Figure 4. After linking from the Argus index record in Trove to a particular article

Option 2. Add index annotations (not tags or comments) to the digitised article

Index terms would be added to the article, preserving the structure of the subheadings (See fig. 5). Clicking on a term could return other articles with the same index term. (But not articles tagged with the same term?) Users could navigate to the (chronologically) next article with the same index term.

Normal users couldn't add or edit index terms. Maybe [redacted] and [redacted], and perhaps some other users, could have special permissions to be able to add/edit/remove index terms?

Will users be confused as to what these are, why they are not editable, or why they appear only for such a selected subset of data?

Index terms would not be included in the tag cloud.

This solution depends on being able to automatically identify the correct article.

The screenshot shows the Trove website interface. At the top, the Trove logo is displayed with the tagline "Digitised newspapers and more". Navigation links for "Home", "About", "Browse", and "Help" are visible. A search bar and a "Sign up" button are also present. The main content area displays a newspaper article from "The Argus (Melbourne, Vic. : 1848 - 1956)" dated "Monday 29 August 1870". The article is titled "TELEGRAPHIC DESPATCHES." and contains two main sections: "QUEENSCLIFF, SATURDAY." and "BELFAST, SATURDAY." The article text describes the Pharos ship towing the stern port quarter and 50 feet of the deck of the wrecked barque Sussex into Swan Bay, and Mr. P. J. Martin's address at a meeting of the electors of the Western Province. On the left side of the article, there is a sidebar with several sections: "Lists (None yet)", "Tagged (None yet)", "Index Terms" (listing "Warrnambool" and "Weather"), "Comments (None yet)", and "Electronically Translated Text". The "Index Terms" section is highlighted, showing terms like "Warrnambool" and "Weather" with corresponding article snippets.

Figure 5. An article displaying index terms (in a similar way to tags).

Option 3. Add index data to the digitised article as tags

Index terms would be added to the article as tags (see fig. 6). Clicking on a term would return other articles with the same term or tag.

Other users couldn't edit or remove the terms, but could add new tags. [redacted] and [redacted] could remove or edit them using a particular (shared) user account.

The structure of sub-headings would not be preserved. Mark-up of proper names would be lost. Users could not navigate to the (chronologically) next article with the same index term except by clicking on the tag and sorting the results by date.

This solution relies on being able to identify the article the index refers to.

The screenshot shows the Trove website interface. At the top, the Trove logo is displayed with the tagline "Digitised newspapers and more". Navigation links for Home, About, Browse, and Help are visible. A search bar and a "Sign up" button are also present. The main content area displays a newspaper article from "The Argus (Melbourne, Vic. : 1848 - 1956)" dated Monday 29 August 1870. The article is titled "TELEGRAPHIC DESPACHES." and contains two main sections: "QUEENSCLIFF, SATURDAY." and "BELFAST, SATURDAY." The article text describes the towing of a wrecked barque and a meeting of the electors of the Western Province. On the left side of the article, there is a sidebar with various interactive elements: "Lists (None yet)", "Tagged" (with tags like "Warrnambool", "Weather", "laying of foundation stone at new masonic hall", "heavy gale at sydney"), "Comments (None yet)", and "Electronically Translated Text". The article text is displayed in a serif font, and the sidebar elements are in a sans-serif font.

Figure 6. Article with index terms added as tags

Index terms would be included in the tag cloud, which would provide some kind of browse option. To see only Argus index terms, users would need to either:

- search for the word “Argus” (And all tags would need to have the prefix “Argus”) (See fig 7.) OR
- view the “ArgusIndex” user profile (See fig. 8)

The Trove tag cloud can display up to 3,000 tags at any one time. The tag cloud can also be searched. The Argus index contains 174,323 different index terms at the moment.

The screenshot shows the Trove website interface. At the top, there is a navigation bar with links for Home, About, Site news, Statistics, Tags, Help, Forum, and Contact us. Below this is a search bar with a 'Search' button and options for 'Available online', 'Australian content', and 'In my libraries'. A menu bar lists various content types: Books, Pictures, photos, objects, Journals, articles and data sets, Digitised newspapers and more, Music, sound and video, Maps, Diaries, letters, archives, Archived websites (1996 - now), People and organisations, and Lists.

The main content area is titled 'Tags' and includes 'Display options' with filters for 'Containing:' (set to 'Argus'), 'Display only:' (set to '100 top tags'), 'Added:' (set to 'anytime'), and 'By:' (set to 'anyone'). Below the filters is a large tag cloud where the word 'ARGUS' is the most prominent, followed by other terms like 'ARGUS - Abolition', 'ARGUS - Ballarat', 'ARGUS - Banalla', 'ARGUS - Birds', 'ARGUS - Cake', 'ARGUS - Carlton', 'ARGUS - Cattle', 'ARGUS - Charities', 'ARGUS - Child abuse', 'ARGUS - Chinese', 'ARGUS - Church of England', 'ARGUS - Death', 'ARGUS - Diners and dining', 'ARGUS - Diphtheria', 'ARGUS - Domain gardens', 'ARGUS - Drowning', 'ARGUS - Education', 'ARGUS - Education act', 'ARGUS - Education Department', 'ARGUS - Elections', 'ARGUS - Emerald Hill', 'ARGUS - Entertainers', 'ARGUS - Factory work', 'ARGUS - Finance', 'ARGUS - Fire control', 'ARGUS - Fires', 'ARGUS - Fisheries', 'ARGUS - Fishing', 'ARGUS - Fitzroy', 'ARGUS - France', 'ARGUS - Gambling', 'ARGUS - Geelong', 'ARGUS - Gold', 'ARGUS - Government House', 'ARGUS - Great Britain', 'ARGUS - Harrison, James', 'ARGUS - Health, Central board of', 'ARGUS - Higinbotham, George', 'ARGUS - Holidays', 'ARGUS - Homeless', 'ARGUS - Homicide', 'ARGUS - Immigration', 'ARGUS - Industrial accidents', 'ARGUS - Insecticide', 'ARGUS - Intercolonial relations', 'ARGUS - Judiciary', 'ARGUS - Kew', 'ARGUS - Labour', 'ARGUS - Land', 'ARGUS - Land acts', 'ARGUS - Land and Works, department of', 'ARGUS - Land selection', 'ARGUS - Larrikin', 'ARGUS - Law', 'ARGUS - Legal profession', 'ARGUS - Licensing', 'ARGUS - Litigation', 'ARGUS - Local government', 'ARGUS - Maintenance (family)', 'ARGUS - Marriage', 'ARGUS - Meat preservation', 'ARGUS - Medical profession', 'ARGUS - Melbourne', 'ARGUS - National Agricultural Society of Victoria', 'ARGUS - New South Wales', 'ARGUS - Pacific Islands', 'ARGUS - Parks and gardens', 'ARGUS - Parliaments', 'ARGUS - Pastoral Industry', 'ARGUS - Patents and Inventions', 'ARGUS - Railway accidents', 'ARGUS - Railways', 'ARGUS - St Kilda', 'ARGUS - Sandhurst', 'ARGUS - Schools', 'ARGUS - Singers', 'ARGUS - Tariffs', 'ARGUS - Teaching', 'ARGUS - Telegraph', 'ARGUS - United States', 'ARGUS - Vandalism', 'ARGUS - Wages', 'ARGUS - Weather', 'ARGUS - Wetlands', and 'ARGUS - Yarra River'.

At the bottom of the page, there is a 'Bookmark this tag cloud' link and a footer for the National Library of Australia, including a disclaimer, privacy policy, terms of use, copyright information, and version number (4.1.1).

Figure 7. Index terms in tag cloud, with “ARGUS” prefix

one search... A wealth of information

All Books Pictures, photos, objects Journals, articles and data sets Digitised newspapers and more Music, sound and video Maps Diaries, letters, archives Archived websites (1996 – now) People and organisations Lists

Available online Australian content In my libraries [Advanced search](#) [Search tips](#)

Information about Trove user: ArgusIndex

[View user profile in the Trove forum](#)

[Tags](#) [Recent comments](#) [Text corrections](#) [Recent merge/splits](#) [Lists](#)

Display options

Containing: Display only: Added: By:

[top tags](#)

Abbotter Abuse Language **Accidents** Acclimatisation society **Actors** Alcohol abuse **Ballarat** Banalla Birds Cabs **Carlton** Cattle charities
 Child abuse **Chinese** Church of England Death Diners and dining Diptheria Domain gardens Drowning Education Education act **Education**
Department Elections Emerald Hill **Entertainers** Factory work Finance Fire control **Fires** Fisheries **Fishing** Fitzroy France Gambling **Geelong**
Gold Government House Great Britain Harrison, James **Health, Central board of** Higinbotham, George Holidays homeless **Homicide**
 Immigration industrial accidents Inanicide Intercolonial relations Judiciary Kew **Labour** Land Land acts Land and Works, department of
 Land selection Larrikin **Law** Legal profession Licensing **Litigation** Local government Maintenance (family) **Marriage** Meat
preservation Medical profession **Melbourne** National Agricultural Society of Victoria **New South Wales** Pacific Islands Parks and gardens
Parliaments Pastoral Industry **Patents and inventions** Railway accidents **Railways** St Kilda Sandhurst schools Singers Tariffs **Teaching**
 Telegraph United States Vandalism **Wages** Weather wetlands Yarra River

[Bookmark this tag cloud](#)



[Disclaimer](#) [Privacy](#) [Terms of Use](#) [Copyright](#) Version 4.1.1

Figure 8. Index terms in tag cloud for user "ArgusIndex"

Trove tags do not allow punctuation other than apostrophe, hyphen, parentheses, comma and forward slash.

There is a 60 character limit on tags in Trove. In order to add new terms or edit existing ones, either:

- the limit would need to be increased for all users
- the limit would need to be increased for the ArgusIndex user
- indexing practice would need to be adjusted, by breaking index phrases into multiple tags. This may also achieve a greater hit rate of related articles.

For example, in future if trying to add the index term "Melbourne Public Library displays Denkmuler aus Aegypten und Aethiopiien by Richard Lepsius, Monumens Egyptiens by C. Leemans, and Holy Land, Syria, Idumea, Arabia, Egypt and Nubia by David Robert" as a tag, it could be broken up in the following way:

- Melbourne public library
- Denkmuler aus Aegypten und Aethiopiien by Richard Lepsius
- Monumens Egyptiens by C Leemans
- Holy Land, Syria, Idumea, Arabia, Egypt, Nubia by D Robert

An article with these tags would not be related to other articles tagged with the full phrase.

Option 4. Create a specialised interface for viewing the index, powered by Trove

The Argus index would still be available as a separate service, and is likely to retain similar functionality to the current service. This is the most labour intensive solution. (See fig. 9)

NATIONAL LIBRARY OF AUSTRALIA

The Argus.

HOME • BROWSE HEADINGS • BROWSE INDEX • SEARCH • HELP • CONTACT

Powered by Trove

THE ARGUS INDEX

1870 TO 1879

Geraldine Suter, Chief Indexer
John Hirst, Editor-in-Chief

© 2005

The Argus, published in Melbourne from 1846-1957, was the quality newspaper of record in the nineteenth and first part of the twentieth century. This index to the Argus covers the decade 1870-1879.

Think of this website as consisting of a core of ten separate and linked annual indexes. The indexes can be:

- browsed directly with the [Browse Index](#) function
- accessed through a cumulated list of headings for all years with [Browse Headings](#)
- approached through the [Search](#) function

The site does not provide online access to the full text of the paper. A separate page describes how to [get copies of articles](#) found through the index.

HOME • BROWSE HEADINGS • BROWSE INDEX • SEARCH • HELP • CONTACT

THE ARGUS INDEX

Figure 9. A specialised interface for the Argus Index, powered by Trove

Option 5. In addition to Option 1, add an extra "Argus index" zone, when viewing newspaper articles

When a user searches the newspapers zone in Trove, show them any hits on the Argus Index at the top right of the search results page (fig. 10). The zone wouldn't appear when looking at any other results screens.

Don't show the extra zone on any of the other results pages.

The screenshot shows a search results page for 'John Olsen' on the Trove website. The search bar at the top contains 'john olsen' and a 'Search' button. Below the search bar are filters for 'Available online', 'Australian content', and 'In my libraries'. The main results area is titled 'Digitised newspapers and more' and shows a list of search results. The first result is 'COURT JOHN OLSEN, I.O.F.' from 'The Brisbane Courier (Qld. : 1864 - 1933) Monday 16 December 1929 p 13 Article'. The second result is 'JOHN OLSEN MEMORIAL.' from 'The Brisbane Courier (Qld. : 1864 - 1933) Tuesday 2 November 1926 p 3 Article'. The third result is 'THE LATE MR. JOHN OLSEN.' from 'The Brisbane Courier (Qld. : 1864 - 1933) Friday 29 October 1926 p 3 Article'. The fourth result is 'MEMORIAL TABLET. LATE MR. JOHN OLSEN.' from 'The Brisbane Courier (Qld. : 1864 - 1933) Thursday 4 November 1926 p 20 Article'. The fifth result is 'JOHN OLSEN'S DEATH. THE INQUEST. Perth, March 12.' from 'Kalgoorlie Western Argus (WA : 1886 - 1916) Tuesday 19 March 1907 p 36 Article'. On the right side of the page, there is a sidebar titled 'Argus Index Results' which lists key dates: 'Olsen, John - obituary 14 Apr 1871', 'Olsen, John - fallen ill 9 Mar 1871', 'Olsen, John - never felt better 7 Mar 1871', 'Olsen, John Jnr - forging and uttering 18 May 1871', and 'Olsen, John - promotion to Chief Constable 30 June 1871'. Below this sidebar is a 'Books' section with two results: 'John Olsen / Deborah Hart' and 'John Olsen / by Virginia Spate'. The page also includes a left sidebar with various filters and a bottom section with 'more' and 'Books' links.

Figure 10. An extra "Argus Index" zone, only shown on the "Newspapers" results screen.

Options for the future of the Argus Index (take 2)

Last Edited: 30 January 2012 by J. Meakins

Table of Contents

Summary and Estimates.....	2
Option A. Add index terms to the digitised article, an interface to manage the terms, and a browse interface (was option 2)	3
Overview.....	3
Benefits	3
Drawbacks.....	3
Possible issues.....	3
Detailed description and mock-ups.....	3
1. View index terms on an article	4
2. Load index terms	5
3. Edit index terms.....	6
4. User management	8
5. Browse	10
Option B. Publish the index as a PDF, searchable in Trove ...	13
Overview.....	13
Tasks	13
Benefits	14
Drawbacks.....	14
Possible issues.....	14
Detailed description and mock-ups.....	15
Option C. Retain the existing stand-alone application.....	18
Overview.....	18
Tasks	18
Benefits	18
Drawbacks.....	18
Possible issues.....	18
Detailed description and mockups	19

Includes

estimates 31/1/12

Summary and Estimates

Feature	Benefits	Challenges/Drawbacks
<p>A. Add index terms to the digitised article, + an interface to manage the terms + a browse interface</p>	<ul style="list-style-type: none"> - Preserves the structure and ability to browse the index - Index terms are added to the article, making them more discoverable - No ongoing data load - May facilitate inclusion of other indexes in Trove 	<ul style="list-style-type: none"> - Linking of index references to the correct article may not be accurate enough - Unclear whether this will fully cater to other indexes (further investigation needed) - Significant amount of work - Browse of index terms does not include the ability to search
<p>B. Publish the index as a PDF, searchable in Trove</p>	<ul style="list-style-type: none"> - Preserves the structure and ability to browse the index - The index is searchable in Trove - If data is supplied as a PDF in future, ongoing data load is easy 	<ul style="list-style-type: none"> - Loading a PDF to DCM is possible but somewhat inelegant and unknown (but the PDF could be hosted elsewhere.) - Index terms are not located with the digitised articles - Div 1 are stakeholders but are yet to be consulted - Will indexers be able to deliver a PDF in future, including links? - Some links will lead to the wrong page, due to differences in the Argus version indexed and digitised
<p>C. Retain the existing stand-alone application</p>	<ul style="list-style-type: none"> - Preserves the structure and ability to browse the index 	<ul style="list-style-type: none"> - The index is not searchable or visible in Trove and so the value is lost to Trove users. - Users have to search Trove for the digitised article - Index terms are not located with the digitised articles - A separate application must continue to be maintained - Data must continue to be converted and loaded

TODO – estimates

Option A. Add index terms to the digitised article, an interface to manage the terms, and a browse interface (was option 2)

Overview

Index terms for the 1870s and 1880s would be automatically added to the digitised article in Trove, preserving the structure of the subheadings and navigation to other articles with the same index. The full list of index terms could be browsed.

New index terms for other years will be added and edited directly on the Trove interface, by users with special permissions. Some users may possibly be able to manage other user accounts.

This solution depends on being able to automatically identify the correct article.

Benefits

1. The structure and ability to browse the index is preserved
2. Index terms are added to the article, adding valuable keywords and making articles more discoverable
3. This solution may be able to be used by other indexing projects, and may inspire new indexing projects
4. Index terms can be added directly to the articles in Trove, there will be no ongoing data load for this index

Drawbacks

1. This option will take a significant amount of time to implement

Possible issues

1. Linking to the digitised article will not be completely accurate. It is still unknown whether an acceptable level of accuracy can be met.
2. The Trove interface will need to explain to users that only the Argus Index is included at the moment, and that not all years have been indexed.
3. Index terms are unlikely to be compatible with projects such as the “climate change” research project, as those users are covering a broad range of titles, whereas index projects tend to take ownership of a particular title, and prefer not to allow other users to add index terms to that title.
4. If other indexes will be incorporated into Trove in future, investigation into the structure and requirements of those indexes is needed.
5. This solution has not yet been discussed with the Argus Index owner and may not suit her requirements, or may not be used in the way we expect (index terms added continuously, by many users).
6. There may need to be a way to search all index headings and to search sub-headings of a particular index heading. This functionality is currently available through the existing Argus Index Online.

Detailed description and mock-ups

The following sections are included below:

1. View index terms on an article
2. Load index terms
3. Edit index terms
4. User management
5. Browse

2

1. View index terms on an article

Index terms would be loaded to the digitised article, preserving the structure of the subheadings. The index terms will be similar to tags but will be more controlled and structured. Index terms will not be included in the tag cloud. They will not be editable by normal users. See Figure 1 for an example of how index terms may appear.

Navigation will be provided to the next and previous article with the index term (in chronological order). Clicking on an index term will show a list of all articles with that index term.

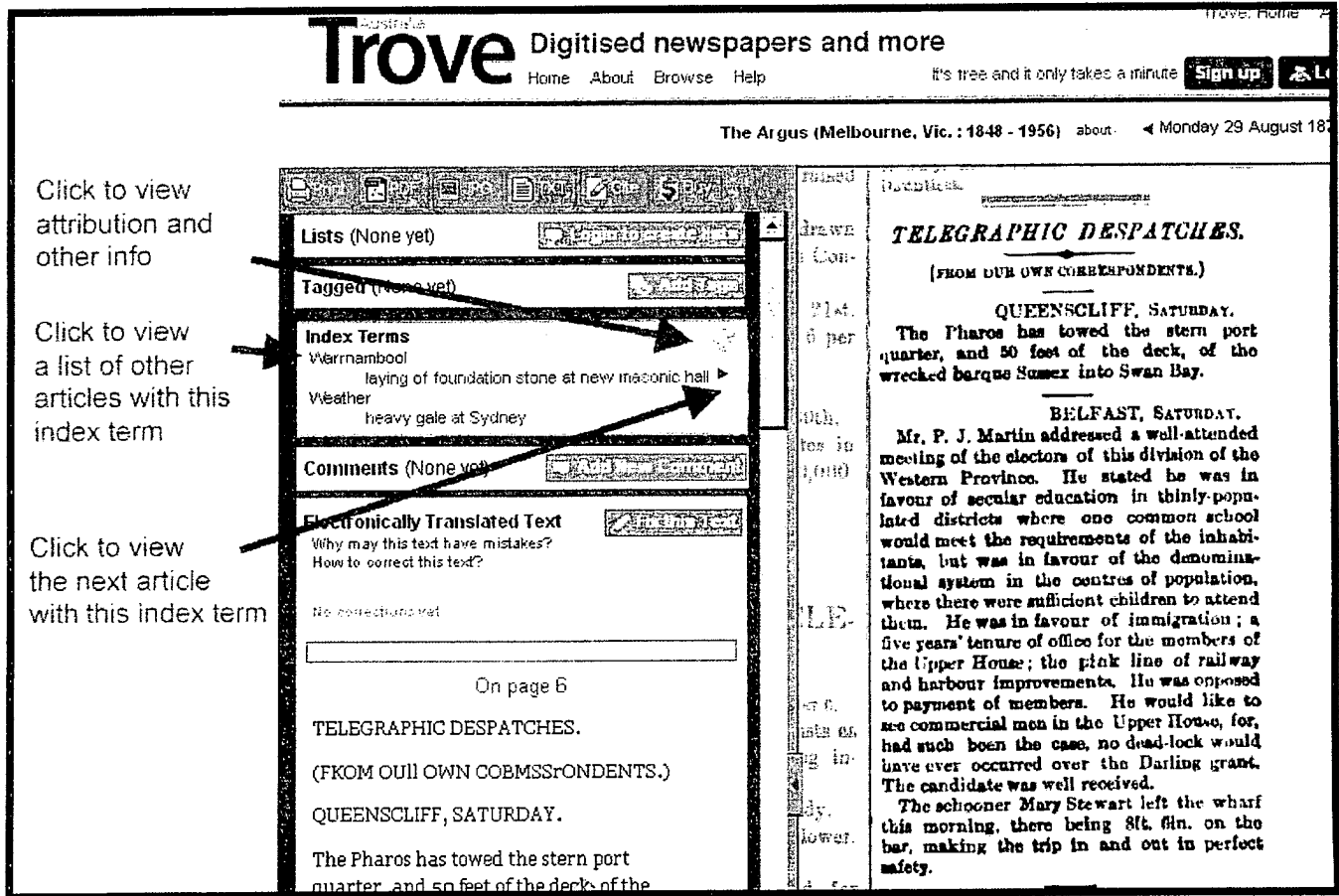


Figure 1. An article displaying index terms (in a similar way to tags).

Tasks

Initially

- Write an information and attribution page for the Argus Index describing scope and coverage (linked from both the Argus title page and articles with index terms)
- Update the Trove API to include index terms with article metadata 1d
- Update the Trove relevance-ranking algorithm to incorporate index terms

On the existing "View article" screen:

- o Interface design to display the index terms 2d
- o Display the index terms 2d
- o Link index terms to a results page listing all the articles with that index term 0.5d
- o Link to the chronologically next article with the same term 1d

When further data is received for the Argus Index

- No further data will be accepted. Indexers must add terms directly to Trove.

In future

- If other indexes are added in future, we would need to consider whether this interface meets their needs.

7

2. Load index terms

In order to show index terms on the digitised article, article references must be automatically matched to the digitised article. It is still unknown whether an acceptable level of accuracy can be met. There are several challenges which need to be considered.

- The Argus Index refers to the Melbourne edition of the Argus; however the NSW edition has been digitised in Trove. The two editions are not exactly the same, so the references in the Argus Index may not match up to the digitised articles.
- The Argus Index references the page number and column, which could contain more than one article.
- Not all words in the index term are included in the article text, and not all words in the digitised article have been correctly OCR'd.

Automatic matching for the index heading and reference:

abattoirs > increased water supply at city abattoirs 9 Feb 1870 5a [deputation]

might perform a search such as:

TITLE: The Argus (Melbourne) AND

DATES: 09/02/1870 AND

(abattoirs OR increased OR water OR supply OR city)

NOT advertising¹

although the exact query is yet to be developed. The matching software would then examine the results and find the article with the highest number of matching words, and the closest page number (and column) to that listed in the reference.

To manage problems with inaccuracies, a data load report should be produced automatically for the Argus Index owner (Geraldine Suter) to review and action. The report should include:

- o terms which could not be matched to an article
- o terms which were matched to an article with a degree of uncertainty

Tasks

Initially

- ~~update~~ *develop* *develop* *implement* *database structures* *3d*
- *modify* *newspaper* *index* *to support* *index terms* *2d*
- Develop an article matching algorithm *2w*
- Load the existing Argus index data (1870s) and the new data on offer (1880s) into Trove, associating each reference with the correct article, in a way that also enables navigation and browsing. *2w*
- Produce a report of non-matches and uncertain matches *1d* *+ edit* *to be added*
- *review* *update* *the* *Trove* *relevance* *ranking* *algorithm* *+* *newspaper* *index* *term*

When further data is received for the Argus Index

- n/a

$$dev = 270 \times 1.5 = 40.5 = 8w$$

$$BA = 0.5w \cdot 2/3$$

In future

- If other indexes are added in future, this data loading program may not necessarily be able to be reused.

✓ *scobleto*
Have to support
controlled vocab, but
not able to add later
without huge work

http://trove.nla.gov.au/newspaper/result?q=&anyWords=increased+water+supply+city+abattoirs&fromdd=09&frommm=02&fromyyy=1870&todd=09&tomm=02&toyyy=1870&l-title=13&l-category=Article|category%3AArticle&l-category=Detailed+lists%2C+results%2C+guides|category%3ADetailed+lists%2C+results%2C+guides&l-category=Family+Notices|category%3AFamily+Notices&l-category=Literature|category%3ALiterature&l-word=*ignore*|*ignore*

3. Edit index terms

An interface is needed to manage index terms. The interface should provide the following functionality:

- Add a new term
 - in a particular position
- Delete a term
- Change the indent level of a term
- Edit the wording of a term

Only authorised users may access this interface – not normal users. The interface could be built into the existing article interface, or as a separate screen. Further interface design is still required. See Figure 2 for an idea of how this interface might appear to Trove users with appropriate permissions.

If the index terms needed to be limited to a controlled vocabulary, further complexity would be added. This has not been considered in the scoping of this task. It is considered briefly below under Exclusions.

The screenshot shows the Trove website interface. At the top, the Trove logo is displayed with the tagline "Digitised newspapers and more". Navigation links include Home, About, Browse, and Help. There are also links for "Sign up" and "Login". The main content area shows a newspaper article from "The Argus (Melbourne, Vic. : 1848 - 1956)" dated "Monday 29 August 1870". The article title is "TELEGRAPHIC DESPATCHES." and it includes sub-sections for "QUEENSCLIFF, SATURDAY." and "BELFAST, SATURDAY." The sidebar on the left contains several sections: "Lists (None yet)", "Tagged (None yet)", "Index Terms" (with sub-items like "Warrnambool" and "Weather"), "Comments (None yet)", and "Electronically Translated Text".

Figure 2. An interface is needed to manage index terms

Exclusions

1. Inclusion of a controlled vocabulary. If this was added in future, questions to consider include:
 - a. Would the vocabulary terms be a suggestion or a limit?
 - b. Can the vocabulary be added to / modified?
 - c. Would the vocabulary apply to the Argus Index only, or all indexes?
 - d. Would a change to a term in the vocabulary retrospectively change all occurrences of that term?
2. Bulk changes to index terms. *or merge index terms*
3. A way for senior users to review index terms added by normal users.

~~4. Merge of index terms~~

(15-20d.)
matea: 40 x 15 = 6w
dev 3w x 15 = 45w

Tasks

Initially

- ~~Interface design for an edit interface~~ - use string
- ~~Implementation of the edit interface~~ ~~3w~~ - 3w

← 1000 KLW

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future, we would need to consider whether this interface meets their needs.

4. User management

User management requirements have not yet been discussed with the Argus Index manager (Geraldine Suter). The assumption has been made that index projects will take ownership of a particular title (or possibly a set of titles if the periodical has ceased and continued under a different name). People outside the project must not be able to add or change index terms on those titles, as the projects use a very controlled vocabulary of index terms.

Two possible user management models are described below. It is not yet known which of these models is most appropriate for the Argus Index project, or whether these models will be appropriate for indexes incorporated into Trove in the future. It is possible that the Argus Index owner will not want other users to add index headings directly to Trove.

Model 1 - Flat structure, administered by NLA IT staff

Indexers

- Can add, edit and delete any index terms on a single title
- Are granted this role by NLA IT CAB staff

AA

This model might be suitable for the Argus Index as only one title exists in Trove (not multiple titles). ~~██████████~~ currently reviews all indexing activity, so may prefer to be the only user adding the index terms to Trove, while the volunteer indexers continue to work in a paper based format.

This model is only suitable for a small group of users or a group of users that does not change regularly; otherwise it would require too much ongoing work for NLA IT CAB staff.

Tasks

Initially

- Require user permissions to access the index term edit interface *2d*
- Design and storage of permissions in the user management system *0.5d*

2.5d x 1.5 = 3.75d

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future, we would need to consider whether this model meets their needs.
- *When new users need to be added 30 min*

AA

Model 2 - Two tier, administered by index managers.

Index managers

- Manage a set of titles. For these titles, they can:
 - Add, edit and delete any index terms
 - Make other users *indexers* (and take this permission away)
 - Make other users *index managers* (and take this permission away)

Indexers

- Are assigned to the same set of titles as their index manager.
- Can add, edit and delete any index terms (on the titles they are assigned to)

An interface is required for index managers to give permissions to other users. This could be part of Trove, a standalone interface, or part of the pending Trove admin module. Figure 3 shows an example interface.

This model is suitable for a medium to large group of users, as it allows users to manage themselves.

Indexer user management

Titles

A list of titles for which you and these users can edit index terms.
 - The Argus (Melbourne, Vic. : 1848 - 1956)

Add

Username:

e.g. johnsmith

Level:

Indexer

Add

Remove

Username: avidtrover (Indexer)

Remove

Figure 3. User management interface for index managers

Tasks

Initially

- Require user permissions to access the index term edit interface 2d
- Design and storage of permissions in the user management system 0.5d
- Interface design of user management interface 3d
- Implementation of user management interface 3d

dev: 5.5d x 1.5 = 8
 int: 3.0 x 1.5 = 4.5d =

TODD
 Ash KW

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future, we would need to consider whether this model meets their needs.

5. Browse

The Argus Index Online is primarily a browse-based interface. This allows users to get an impression of the events, news and topics reported on, in context. A browse interface for index terms could be added to Trove by listing index terms on the browse menu for the newspapers zone (see Figure 4).

The top index terms would be listed, as well as an option to view all terms. If other indexes are added in future, this list could include terms from all indexes, not just the Argus Index.

There would be no change to the Argus title page in Trove. <http://trove.nla.gov.au/ndp/del/title/13>

The screenshot shows the Trove website interface. At the top, there is a navigation bar with 'Home', 'About', 'Browse', and 'Help'. The main content area is divided into two columns: 'Browse Issues' and 'Browse Articles'.

Browse Issues

- by Title**
 - The Argus
 - The Mercury
 - The Courier-Mail
 - The Perth Gazette and Western Australian Journal
 - The South Australian Advertiser
 - The Australian Women's Weekly or Browse by cover (New!)
- by State**
 - National
 - ACT
 - NSW
 - VIC
 - QLD
 - SA
 - WA
 - NT
 - TAS
- by Date**

1803

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

Browse Articles

- by Category**
 - Advertising
 - Detailed lists results guides
 - Family Notices
 - News
- by Tag**
 - LRRSA (11221)
 - poem (5820)
 - BMD (4707)
 - Cricket (4081)
 - RWM (3681)
 - Cemetery of interment unknown (2745)
 - Rookwood Necropolis - interments (2608)
 - Cane (2534)
 - China (2511)
 - animation (2235)
- by Index heading**
 - Litigation
 - Labour
 - Accidents
 - Health - central board of
 - Ballarat
 - Land and works, department of
 - Gold
 - Meat

Handwritten annotations: An arrow points to the 'by Index heading' section. A note says 'May be difficult to generate if multiple ~~index~~ ^{title} are index'.

Figure 4. Index headings could be listed on the browse menu for the newspapers zone

May be difficult to generate
if multiple ~~index~~ ^{title} are index

If users choose to view all terms, the top-level terms will be shown in an alphabetical list. Figure 5 shows a display similar to the “Browse headings” screen² in the current Argus Index Online, including a navigation wheel and paging.

In future, if other indexes are added to Trove, a facet could be used to limit browsing of index terms to a single title only.

Trove: Home About Site n

Trove Digitised newspapers and more
Home About Browse Help It's free and it only takes a minute [Sign up](#) [Login](#)

Browse index headings

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

▼ Title
The Argus (Melbourne)

↑
Title facet could be added if other indexes are included

- Aarons, J. 1872, 1 item
- Aarons, Joseph 1870 - 1879, 49 items
- Aarons, Lewis 1871, 3 items
- Aarons, Solomon (money lender) 1870 - 1879, 9 items
- Abattoirs 1870 - 1879, 189 items
 - See also Echuca Slaughterhouse Company
- Abbotsford 1870 - 1879, 10 items
- Abbotsford Football Club 1878, 1 item
- Abbott, David 1875, 1 item
- Abbott, Henry 1872, 1 item
- Abbott, H.H. 1878, 1 item
- Abbott, Joseph Henry, 1877 - 1879, 2 items
- Abduction, 1870 - 1879, 63 items
- Abe, F.W. (Rev) 1873, 2 items
- A'Beckett, - (Mrs) 1875, 2 items
- A'Beckett, E. F. 1876-1879, 3 items
- A'Beckett, L. J. 1870-1879, 19 items
- A'Beckett, R. 1873, 1 item
- A'Beckett, Thomas T. 1870-1879, 35 items
- A'Beckett, W.A.C., 1871-1877, 13 items
- A'Beckett, W. G. 1876, 1 item
- Abel, Frederick Augustus (Prof.) 1871-1877, 2 items
- Aberfeldie 1872-1876, 2 items
- Aberfeldie Coursing Association 1878, 1 item
- Aberfeldy River 1871, 1 item
- Aberline, Joseph 1874, 1 item
- Abernethy, J. M. (Rev.) 1870-1879, 10 items
- Abington, Fred 1877, 1 item
- Aborigines 1870-1879, 324 items
 - see also Boomerangs, Church of England Mission to the Aborigines, Presbyterian Chinese and Aboriginal Missions

Next >

Figure 5. Index terms can be browsed

Possibly have to choose a title first.

When users click on a particular heading in the list, subheadings and references to articles will be shown (see Figure 6).

- Each article reference would be linked to the digitised article.
- A facet could be used to limit to a particular year (or decade) of references.
- Users could navigate back to the full list of top level index headings.

Again, if other indexes are added to Trove in future, a facet could be used to limit to index terms from a single title only. Article references would then need to indicate which title they refer to.

Trove Digitised newspapers and more
Home About Browse Help It's free and it only takes a minute Sign up Login

[Browse full index](#)

Index heading: Abattoirs

see also Echuca Slaughterhouse Company
Henry Walker, boiling down works nuisance 6 Jan 1870 5b [court case]
relocation of abattoirs 21 Jan 1870 supp 7g
overflow at Richmond abattoirs after storm 25 Jan 1870 supp 1d [meeting]
St Kilda abattoirs 8 Feb 1870 supp 1g [meeting] , 25 May 1870 7g [meeting]
increased water supply at city abattoirs 9 Feb 1870 5a [deputation]
inspector of nuisances at Hawthorn reports on inaction on paving and draining of butchers yards 2
inspector of slaughterhouses at Hawthorn to be vested in sergeant of police 22 Feb 1870 7e [mee
licences for Heidelberg slaughterhouses 22 Feb 1870 7b [meeting]
Mr Kyle appointed Brunswick inspector 26 Feb 1870 supp 1d [meeting]
R. Elworthy, St Kilda butcher, nuisance on Caulfield boiling down premises 26 Feb 1870 5e [court
Watson and Patterson, curers, unlicensed slaughter of pigs 4 Mar 1870 6e [court case]
treatment of cattle awaiting slaughter in yards 17 Mar 1870 5d
drain around abattoirs at Emerald Hill 18 Mar 1870 supp 2b [meeting]
Sergeant Ellis, inspector of abattoirs, summons Ferdinand Hann for fees 18 Mar 1870 supp 1g [co
slaughterhouse problems at Brunswick 18 Mar 1870 supp 2a [meeting]
Sunday slaughtering to be abolished in Brunswick 26 Mar 1870 7b [meeting]
slaughtering licence of David Baker, Hawthorn 6 Apr 1870 7a [court case] , 28 Apr 1870 7c [meetin
Alexander Robertson and John Wagner slaughter in unlicensed premises 8 Apr 1870 7b [court cas
inspector of abattoirs and slaughtering at Williamstown 13 Apr 1870 6e [meeting]
John Gray and William Gray, abattoir lessees, fined for offensive smell 22 Apr 1870 6d [court case]
W. W. Richards, boiling down nuisance 22 Apr 1870 6d [court case]
city abattoir nuisance 6 May 1870 7c [meeting] , 3 Jun 1870 supp 2g [meeting]
Melbourne Corporation meeting on new abattoirs act 17 May 1870 6a
Richard Elworthy, Elsternwick boiling down nuisance cases 18 May 1870 supp 1g [court case] , 20
Braybrook inspector of slaughterhouses issues summons 20 May 1870 7e [court case]

Title
The Argus (Melbourne...
↑
Title facet could be added if other indexes are included

Year
1870 - 1879
1870 (90)
1871 (15)
1872 (18)
1873 (5)
1874 (21)
1875 (6)
1876 (9)
1877 (9)
1878 (8)
1879 (8)

Figure 6. Browsing the "Abattoirs" index heading

Tasks

Initially

- Interface design of multiple browse screens
- Implementation of multiple browse screens
 - o generate list of terms
 - o navigation
 - o decade and year facets
 - o link article references to the digitised article

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future:
 - o add a title facet
 - o add newspaper title name to the article reference
 - o consider whether this model meets the needs of the other indexes

Dev - ROMANT Jw r 1.5 = 4.5w
make - 1w x 1.5 = 1.5w

sd → TSDS KW
2.5w w
Ad

Option B. Publish the index as a PDF, searchable in Trove

Overview

The 1860s decade of Argus index data was published as a book³. We could migrate each year of index data currently in the Argus Index Online into an online PDF. The PDF could be hosted by the NLA with the full text searchable in Trove (similar to finding aids and "local history resources" indexed by Trove) and index references linked to the digitised page in Trove.

The layout of the PDF could be styled on the layout of the existing book.

The existing decade (1870s) and the next decade of data on offer (1880s) could be converted to PDF by the NLA. In future, the NLA would only accept the data in PDF format. Digitisation of the 1860s hard copy index is not considered here.

NLA cataloguers would need to create a metadata record in the NLA catalogue for each PDF – one for each year included. (The metadata records will automatically flow through to Libraries Australia and Trove.)

Tasks

Initially

- 0-Design PDF Grant*
1. Convert the existing Argus index data (1870s) and the new data on offer (1880s) into text searchable PDFs – one for each year, with article references linked to the digitised page in Trove. *2w*
 2. Upload the Argus Index PDFs to DCM (or another public location). *1w*
 3. Create metadata records for each Argus Index PDF in the NLA catalogue, including a link to the PDF and a flag to indicate that the PDF should be indexed by Trove. *Non-1*
 4. Organise for data to be provided in PDF format in future.

When further data is received for the Argus Index

1. Upload the Argus Index PDFs to DCM (or another public location). *1/2d*
2. Create metadata records for each Argus Index PDF in the NLA catalogue, including a link to the PDF and a flag to indicate that the PDF should be indexed by Trove

In future

1. An experimental project could be carried out to investigate adding index data as tags on articles. This project could examine the feasibility and usefulness of this approach. It is recommended that this project:
 - a. uses an index that applies to the same edition of the newspaper as that digitised in Trove, to facilitating matching index terms to the correct digitised article
 - b. involves an index owner who is amenable to converting the index to tags
2. The hard copy Argus index from the 1860s and other years could be digitised, OCR'd and made searchable in Trove. References to these indexes can be found here:

Benefits

1. The structure of the index and the ability to browse it is preserved.
2. This solution does not rely on being able to automatically link an index reference with the corresponding digitised article in Trove.
3. The index is searchable in Trove.
4. If data is supplied as a PDF in future, ongoing data load is easy.

Drawbacks

1. The Argus index terms are not added to the digitised article, so do not add additional search terms for users.
2. The Argus Index is searchable in a different zone to the Argus newspaper articles.
3. Some of the page references will link to a page which does not show the article – due to differences in the Argus version indexed and digitised.

Possible issues

1. This solution assumes that a single year of the index is not too large to cause problems in DCM, for the Trove indexer, or for users to open. (If problems arise, the PDFs could be more granular.)
2. The only capability DCM has for uploading PDFs is as a finding aid (PDFs of books are automatically generated from images and do not include text). The Argus Index may need to be uploaded as a kind of finding aid, including a finding aid PI. DCM was modified to handle PDF finding aids in 2011 but this functionality has not yet been used. Alternately, the PDF could be hosted outside of DCM. Also note that the NLA's DLIR project is due to deliver a solution to upload and deliver PDFs with OCR'd text by end 2012.
- ~~3. Will it be difficult to automatically create a searchable PDF from the current Argus Index data?~~
4. Will indexers be able to deliver a PDF in future, including links to the digitised page in Trove?
5. NLA Div 1 are stakeholders but have not yet been consulted.

Detailed description and mock-ups

Users can search Trove for words included in the index, but Trove will only indicate whether their search terms are found in a particular year of the index. The results will appear in the books zone. Users click on the “view online” link to view the PDF of the index. (See Figure 7)

The screenshot shows the Trove website interface. At the top, it says "NATIONAL LIBRARY OF AUSTRALIA" and "Trove". Below this is a navigation bar with categories: All, Books, Pictures, photos, objects, Journals, articles and data sets, Digitised newspapers and more, Music, sound and video, Maps, Diaries, letters, archives, Archived websites (1996 - now), and People and organisations. A search bar contains the text "lightning accident kiandra" and a "Search" button. Below the search bar are filters for "Available online", "Australian content", and "In my libraries", along with links for "Advanced search" and "Search tips".

The main content area is titled "Refine your results: Books". It shows "Showing: 1 - 20 of 8,805" and "Sort by: Relevance". On the left, there are filters for "Format" (Book: 8,805, Braille: 0, Illustrated: 0, Large print: 0, Audio book: 0, Conference Proceedings: 0, Thesis: 0) and "Availability" (Online: 0, Access conditions: 0, Freely available: 0, Unknown: 1). There are also filters for "Decade", "Language", and "Australian Content".

The search results list several books:

- Proceedings relative to the accident by lightning at Heckingham**
[Book: 1783]
At 5 libraries
- Effects of lightning upon the human body, exemplified by a remarkable accident / by Dr. Tilesius**
Tilesius von Telenau, Wilhelm Gottlieb, 1769-1857
[Book: 1824]
Keywords: Electrical injuries; Lightning
At Baillieu Library
- Indexes to the Argus, 1860 / indexed by Geraldine Suter**
Suter, Geraldine
[Book: 1999]
Keywords: Argus (Melbourne, Vic.) - Indexes; Australian newspapers - Indexes; Newspapers - Indexes.
...four killed in lightning accident at Kiandra. 1860-09-20 pg5 col f men
fall down holes at Inglewood 1860-09-24 pg 5 col g...
View online
At National Library

Annotations on the screenshot include:

- An arrow pointing to the "Availability" filter section with the text: "2. Various years of the Argus Index may appear in the list of results".
- An arrow pointing to the "View online" link for the "Indexes to the Argus" result with the text: "3. Click to view the index".

Figure 7. Finding the Argus Index as a result in Trove

Users then search the PDF for the terms they are interested in (see Figure 8). Users can also browse the PDF of the index if they wish.

Accidents
see also Drowning; Fires; Industrial accidents; Mining accidents; Poisons; Railway accidents; Road accidents; Safety; Shipping accidents; Shooting accidents

1860 -04-07 pg 6 col c
driver from Mount Esse station killed in dray accident,
1860-03-02 pg 3 col g
glass globe falls on man in Legislative Assembly,
1860-03-16 pg 5 col b
strange accidents at Theatre Royal, 1860-03-17 pg 5 col c
child swallows seeds of thornapple at Ballarat, 1860-03-21 pg 5 col c

man loses eye in accident with bullock whip, 1860-03-31 pg 5 col g
old man, brought to court at Maryborough on assault charges, injured, 1860-04-04 pg 5 col c
child falls into waterhole at Malmsbury, 1860-04-17 pg 6 col f
boy falls down hole at quoits match, 1860-05-11 pg 5 col g
man injured in earthfall at Prahran railway works,
1860-05-16 pg 5 col a
man dies at Bendigo after falling under load of wood,
1860-06-04 pg 4 col d
woman dies from asphyxiation while drunk, 1860-06-13 pg 5 col g
Buninyong man rescued by Chinese three days after falling into hole 20 ft deep, 1860-06-16 pg 5 col d
part of cornice of new Bank of Victoria building at Hamilton falls to ground, 1860-06-19 pg 5 col c
man killed in blasting operations at Kyneton, 1869-06-29 pg 4 col b
sailor dies due to lack of lighting on Sandridge railway pier, 1860-08-20 pg 6 col e
child suffocates on piece of meat, 1860-08-21 pg 5 col b
child dies after falling down shaft at Inglewood, 1860-09-17 pg 5 col c

drunken man in accident at Theatre Royal, 1860-09-18 pg 5 col d
four killed in lightning accident at Kiandra, 1860-09-20 pg 5 col f
men fall down holes at Inglewood, 1860-09-24 pg 5 col g

Brown, John, Yackanda over stump, 600615 7
Burke, Patrick, killed in
Burridge, -, dies after f
600917 5c
Butt, George, attorney's springcart in Little C
Campbell, -, mounted c horse at Morses Cree
Caulivan, Michael, run
Chaplin, George, falls f
Christopher, James, wo powder, 601121 5a
Clarke, G., Maldon, ac 601106 5d
Coard, -, killed in hors
Coates, -, dies after fal
Cooke, William, kicked
Cox, -, thrown from ca
Davis, Henry, scalped i Creek, 600605 7a
Deuning, Alfred, killed 600702 4e
Donald, John, injured b 600221 5b
Dudden, Alfred Henry, 600818 5c
Earl, Sarah, falls off ca
Easton, James, run ove
Erskin, -, thrown from
Falkner, George, sheph for family, 601022 5
Faris, Frank B., Emeu lark after hunt, 6006
Farrel, Edward, injured

Figure 8. Searching a PDF of the Argus Index

When a user finds a reference to an article they want to read, and click on the page link to return to Trove. The user must then find the article on the page (See Figure 9). In some cases, the article will be located on a different page.

The screenshot shows the Trove website interface. At the top, there is a navigation bar with links for Home, About, Site news, Statistics, Tags, Help, Forum, and Contact us. A search bar contains the word 'kiandra' and has a 'Search' button. Below the search bar, there are options for 'Show search limits' and 'Advanced search'. The main header indicates the document is 'The Argus (Melbourne, Vic. : 1848 - 1956)' and the current page is 'Thursday 20 September 1860', 'Page 5' of 8. There are icons for PDF, PDF (small), and other document types. On the left side, there is a 'Page 5 articles' section with three entries:

- THURSDAY, SEPTEMBER 20, 1860.**
A correspondent, who dates from the Sandhurst Mining Board, has called our attention to what he deems the neglect shown to the gold-producing ...
[Continued from Page 4]
[Article] (466) words.
- TELEGRAPHIC DESPACHES. (FROM OUR OWN CORRESPONDENTS.) SYDNEY, WEDNESDAY.**
Heavy storm took place at Kiandra on Sunday morning. Four men, recently from Sydney, were killed by the lightning. Their names were—John Pateman, Donald Mackinnon, Daniel ...
[Article] (163) words.
- SANDRIDGE DOCKS AND TRAMWAY COMPANY. TO THE EDITOR OF THE ARGUS.**
Sir,—As I am unable to find out what has become of the above-named company, after repeated inquiries, will you kindly give this publicity, that many like myself, who thought this an important ...
[Article] (123) words.
- AGRICULTURAL REPORT. WEDNESDAY, SEPTEMBER 19.**
The produce markets are improving again.

The main content area shows the newspaper page with the title 'THE ARGUS, THURSDAY, SEPTEMBER 20, 1860.' and several columns of text. A 'ZOOM' button is visible at the bottom right of the page preview.

Figure 9. The Argus Index could link to the digitised page in Trove. Users must then find the article.

Option C. Retain the existing stand-alone application

Overview

As the current server is being decommissioned, the existing application could be ported to the new server and infrastructure.

The decade of data currently on offer would be loaded into the existing application. A decision would need to be made as to whether the National Library continues to host the application indefinitely, loading data periodically as it is made available, or not.

Dev ~~18/11~~
1.5 > 1.5 = 84
JFW

Tasks

Initially

1. Port the existing application to the new server and infrastructure 2w
2. Load the new data on offer (1880s) into the existing Argus Index application ~~1w~~ 1.5w
3. On the Argus index Online home page, add a link to the Argus title page in Trove 0.5d
4. ~~Document form or load method~~ 1d (add to run)

When further data is received for the Argus Index

If the NLA decides to continue loading new data:

1. Load the data into the existing Argus Index application 1.5 w

In future

If the NLA decides to continue to host the application:

1. Continue to maintain the Argus Index application

Otherwise:

1. Decide what to do with the Argus Index (again)?

+ Research project to investigate adding index terms as tags

Benefits

1. Preserves the structure of the index
2. Preserves the ability to browse the index

Drawbacks

1. The index is not searchable or visible in Trove
2. Users have to search Trove for the digitised article
3. Index terms are not located with the digitised articles
4. A separate application must continue to be maintained
5. Data must continue to be converted and loaded

Possible issues

1. The existing Argus Index application has no current business owner

enhance to
link to digitised page in Trove?

Detailed description and mockups

The existing Argus Index application will continue to exist (see Figure 10). It includes the ability to:

- search the index terms
- browse the index terms alphabetically
- browse the index terms alphabetically for a particular year

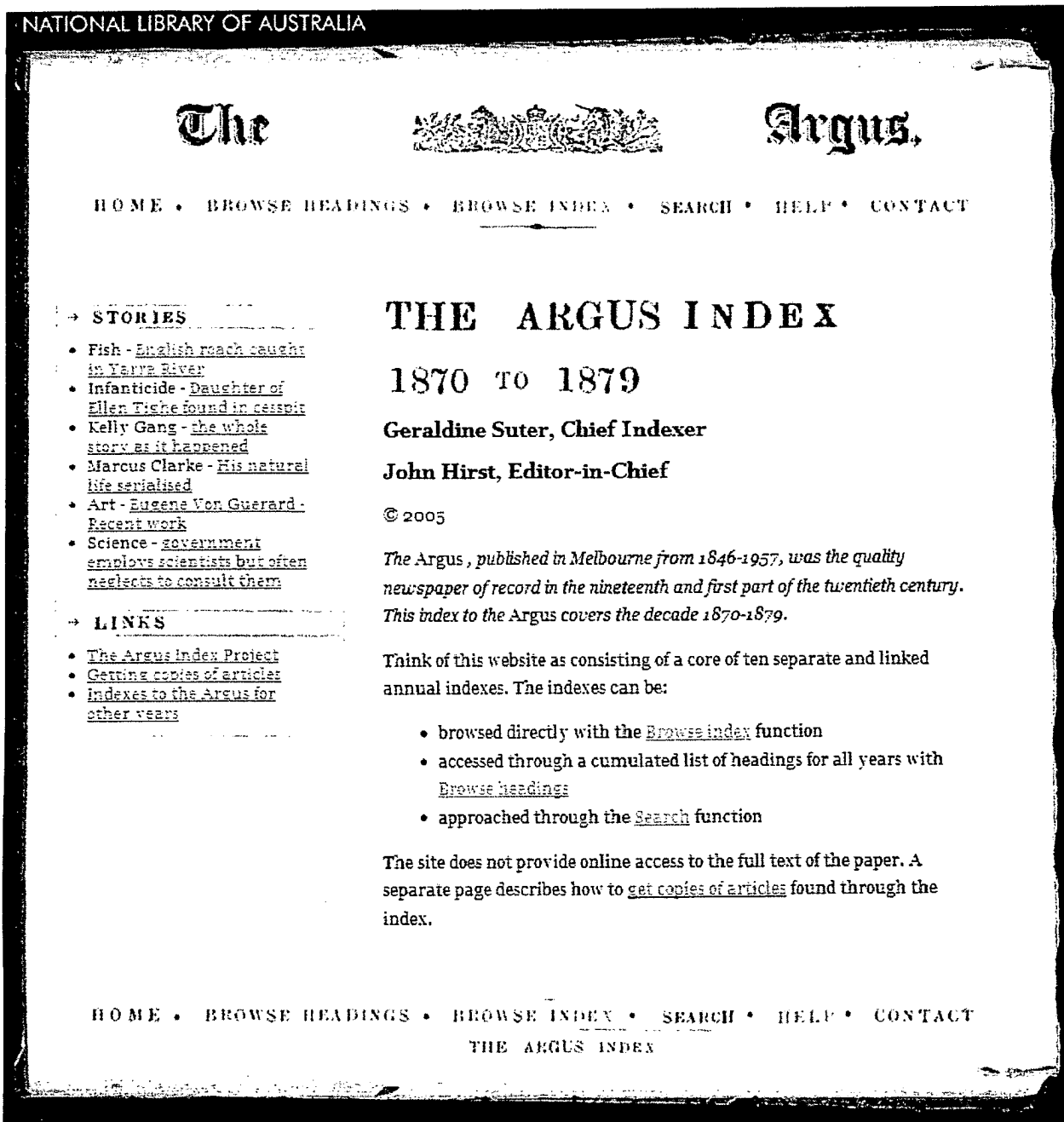


Figure 10. The existing Argus Index application

Options for Argus Indexer communication using the Trove forum

Last Edited: 30 January 2012 by J. Meakins

Table of Contents

Summary.....	Error! Bookmark not defined.
Method 1 – A social group	2
Method 2 - A private forum, which users can apply to join.....	5
Method 3 – A password protected forum	9

Introduction

The Trove forum software could be used for Argus Index volunteers to communicate with one another.

There are several possible ways to do this, such as:

1. A social group
2. A private forum, which users can apply to join
3. A password protected forum

Method 1 – A social group

Benefits:

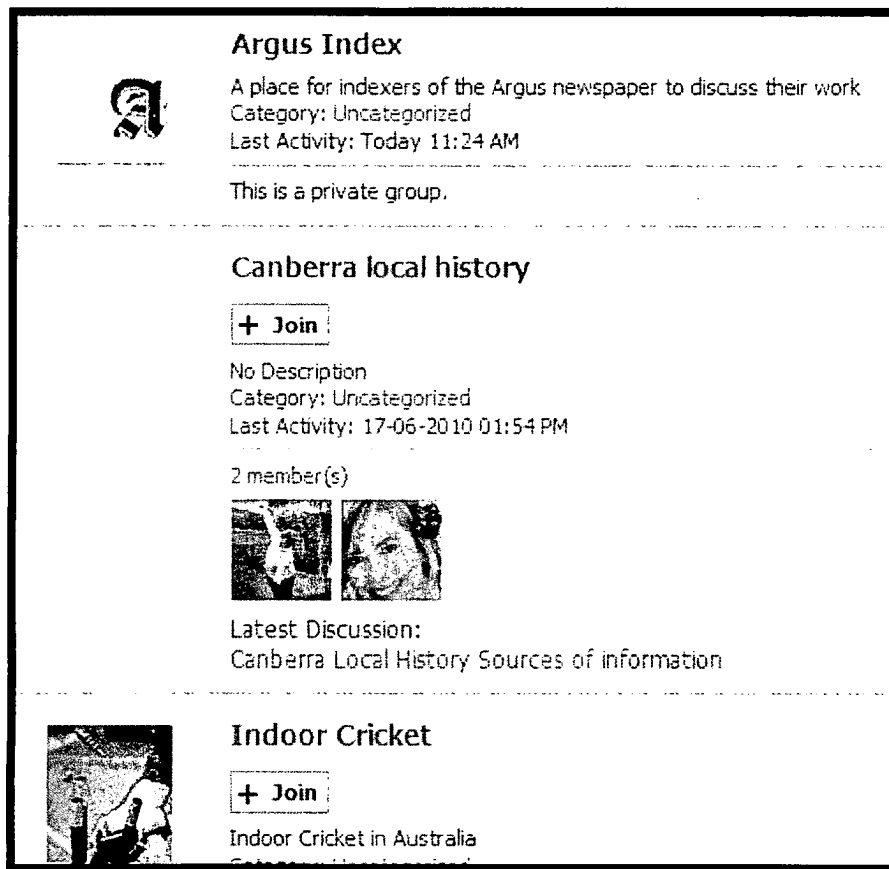
- Only one setting needs to be changed by Div 3 staff.
- Argus Index owner (as the group leader) can manage and moderate the group and its membership relatively easily.
- Users are sent a notification when they are invited to join the group.

Drawbacks:


- Only one group leader can exist and this person can not be changed. (A new group would have to be created instead.)

The Argus Index owner creates an invitation only social group within the forum. She will automatically become the group leader. The group leader can invite other forum members to join the group. The group can then post discussions amongst themselves which are not shown to people outside the group. (See figures 11, 12 and 13)

Division 3 staff may need to provide technical support to the group leader.



Argus Index
A place for indexers of the Argus newspaper to discuss their work
Category: Uncategorized
Last Activity: Today 11:24 AM
This is a private group.

Canberra local history
[+ Join](#)
No Description
Category: Uncategorized
Last Activity: 17-06-2010 01:54 PM
2 member(s)

Latest Discussion:
Canberra Local History Sources of information

Indoor Cricket
[+ Join](#)
Indoor Cricket in Australia

Figure 1. The Argus Index group, in the list of social groups available. The group is invitation only – no “Join” button is shown.

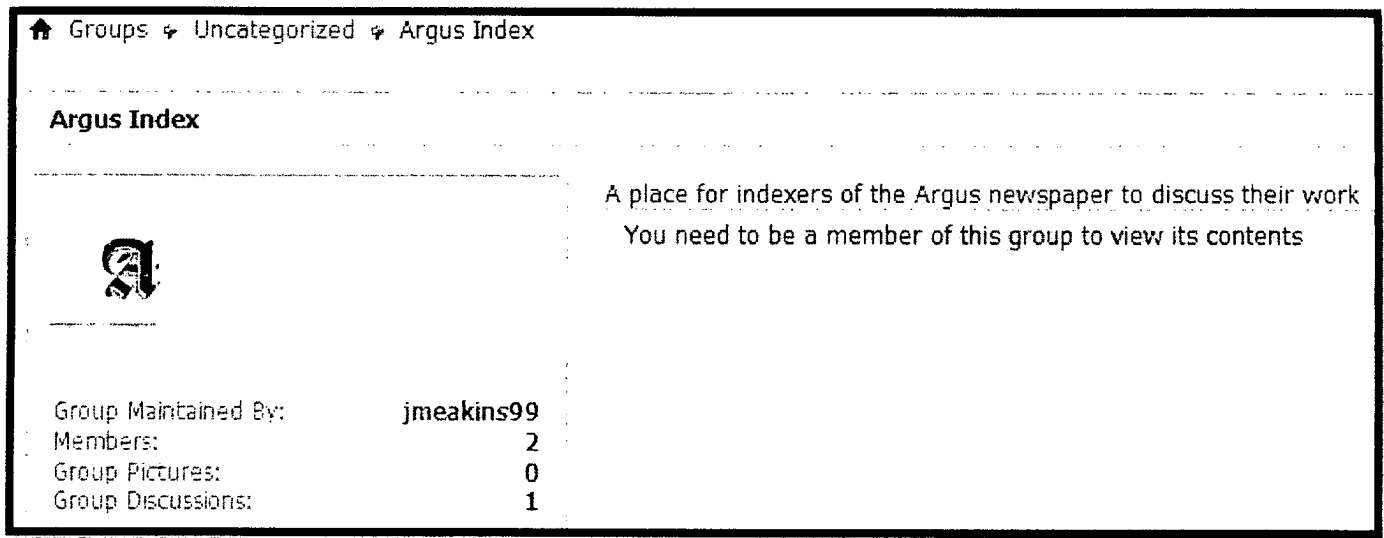


Figure 2. The contents of the Argus Index group can not be seen by non-members

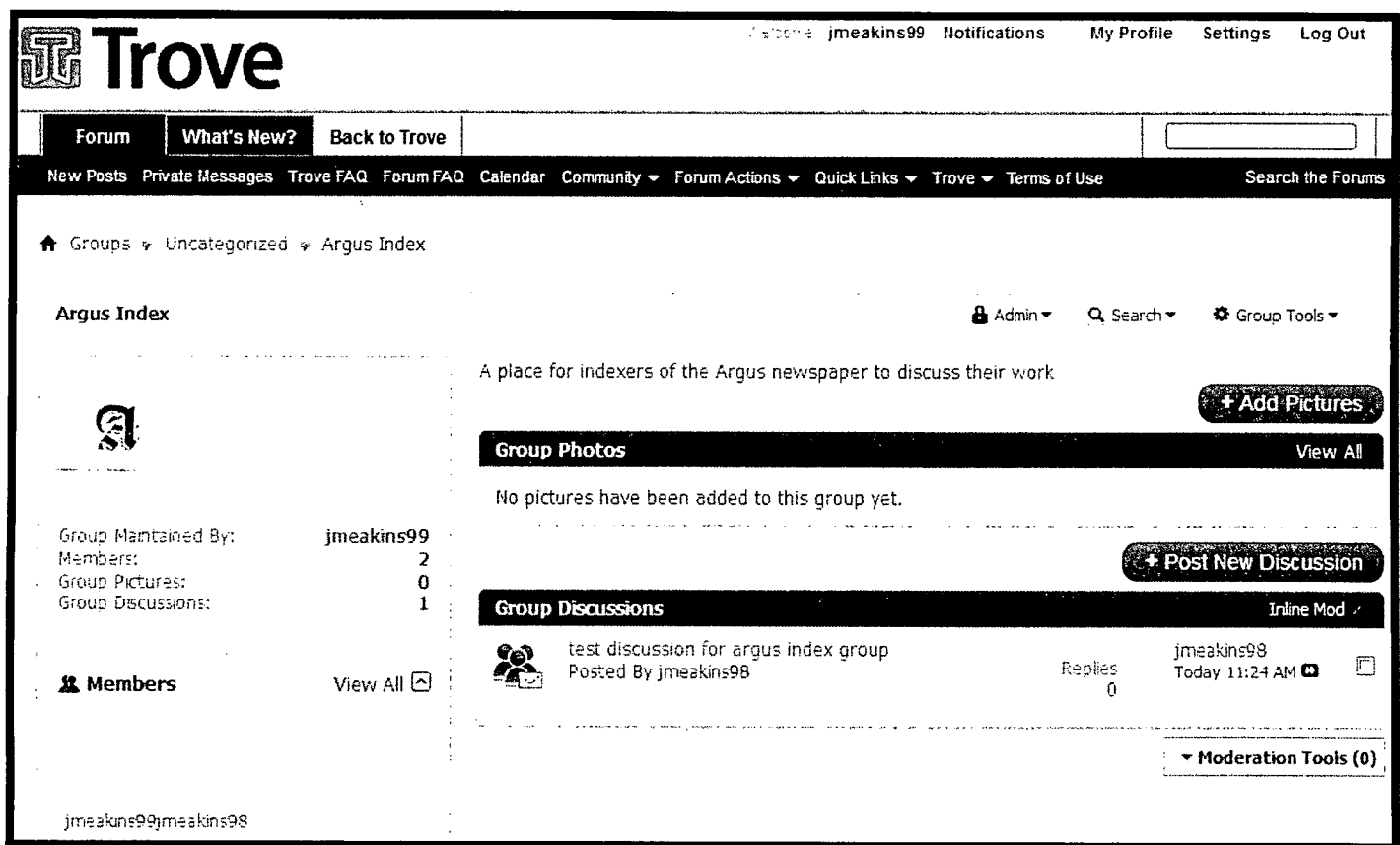


Figure 3. The Argus Index group as it appears to the group leader

To set up an invitation only social group:

1. A Trove forum administrator allows 'invitation only' groups to be created. This only needs to be done once.
 - a) In the forum admin module, choose: "Settings > Options > Group options > Allow join-to-view groups"
2. The group is created by the person who will manage it (the group leader).
 - a) On the Trove forum website, choose: "Community > Groups > Create group"
 - b) Choose the option: "invitation only"
 - c) Choose the option: "users must join to view the content"
3. The group leader invites other users to join the group.
 - a) On the Trove forum website, choose: "Community > Groups > Argus Index" (or whatever the group has been named)

- b) Choose “Admin > Pending & invited members”
- c) Type the username of the person to be invited to join the group and click “Invite”
- d) The invited user will have a new notification next time they log into the Trove forum. They need to read the notification and accept the invitation.

A social group can only have one group leader and this person can not be changed later – it is always the user who first created the group.

Method 2 - A private forum, which users can apply to join

Benefits:

- Argus Index owner (as the group leader) can manage and moderate the group and its membership.
- The forum can have more than one group leader.

Drawbacks:

- A little more complicated than the other methods.
- Some setup need by NLA Div 3 staff.
- The group leader is not automatically notified of new requests to join the group

Users will need to make a request if they want to see the Argus Index forum. The Argus Index manager could act as a “group leader” and approve user requests to join the Argus Index forum.

This option involves the following steps:

NLA Div 3 staff set up the forum and user group:

- Create a “public joinable” user group.
- Assign a group leader to the user group. The group can have more than one group leader, and the group leader can be changed.
- Create a private sub-forum which can only be accessed by that user group.

This process is outlined here: <http://www.cs.umd.edu/faq/forums/creategroup.html>

Argus Indexers ask to join the group:

- Users go to the “Settings > Permission Groups” menu and request to join the user group (see Figure 5).

The Argus Index manager (as the group leader) approves requests to join the group:

- The group leader goes to the “Settings > Permission Groups > Review Join Requests” menu to approve the request (see Figure 6 and Figure 7). The group leader is not notified of new requests.

The private forum will appear as normal to users who have permission to use it (see Figure 4). Other users will not see the forum listed.



















Other			
	Trove News News about Trove, new features, content and planned outages. (Only the Trove administrator may post here)	 	Threads: 0 Posts: 0
	Feedback and Suggestions about the new Trove Forum Do you want changes made to this forum, is it easy to use? Do you need new forum areas set up?	 	Threads: 0 Posts: 0
	Questions Have a question related to your research or Trove? Post it here and see if the Trove community can help you out.	 	Threads: 3 Posts: 5
	Genealogy Are you searching for family history and do want to contact other genealogists?	 	Threads: 0 Posts: 0
	Trove: In and Out of the Classroom Talk about how to use Trove in education.	 	Threads: 0 Posts: 0
	Argus Indexers A place for Argus Indexers to discuss their work	 	Threads: 0 Posts: 0

Figure 4. Those who have access to the private "Argus Indexers" forum will see it listed as usual, with the other available forums.

Related resources:

Restricting access to a forum: <http://www.cs.umd.edu/faq/forums/>

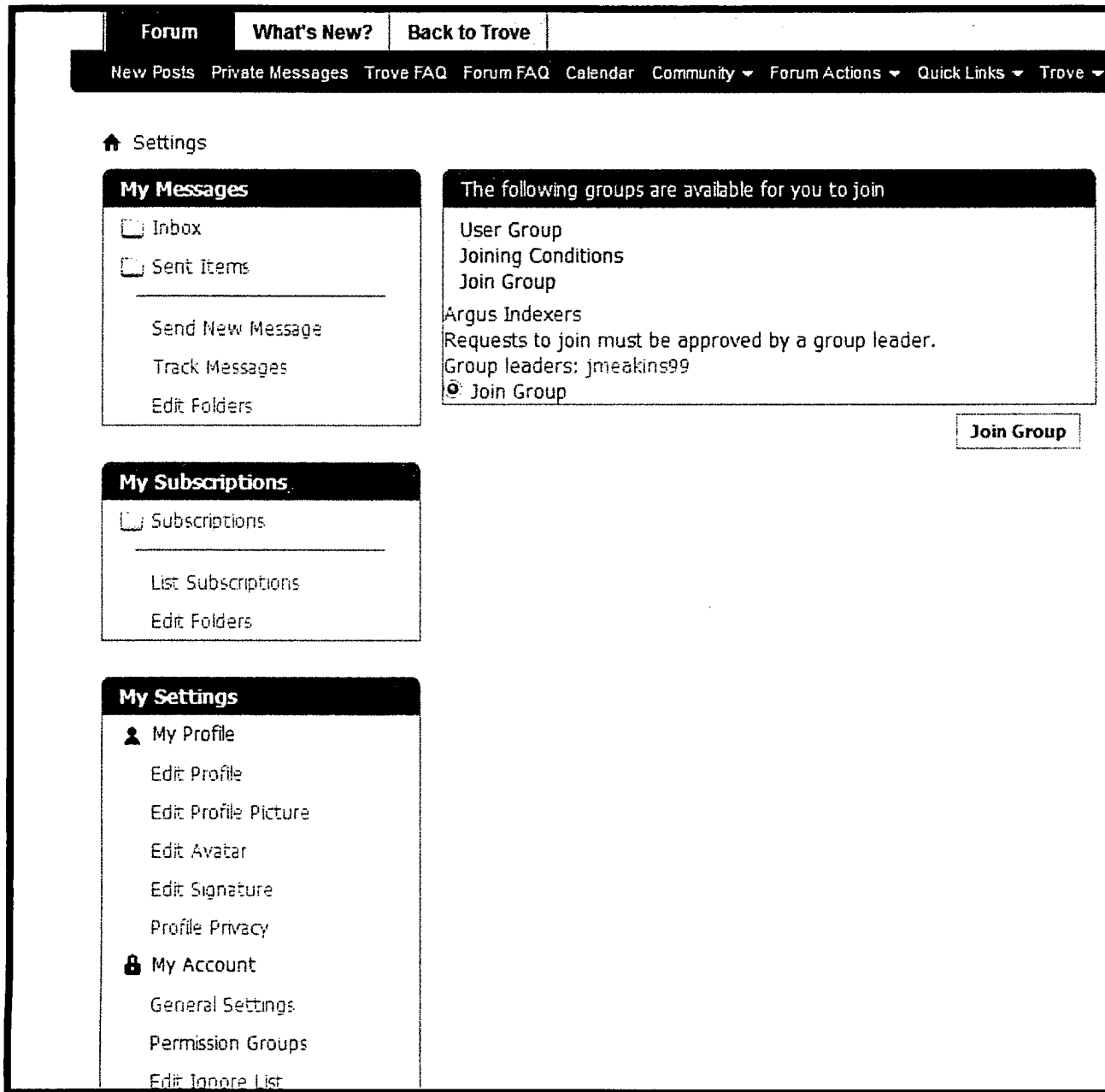


Figure 5. To join a private forum, users go to the "Settings > Permission Groups" menu

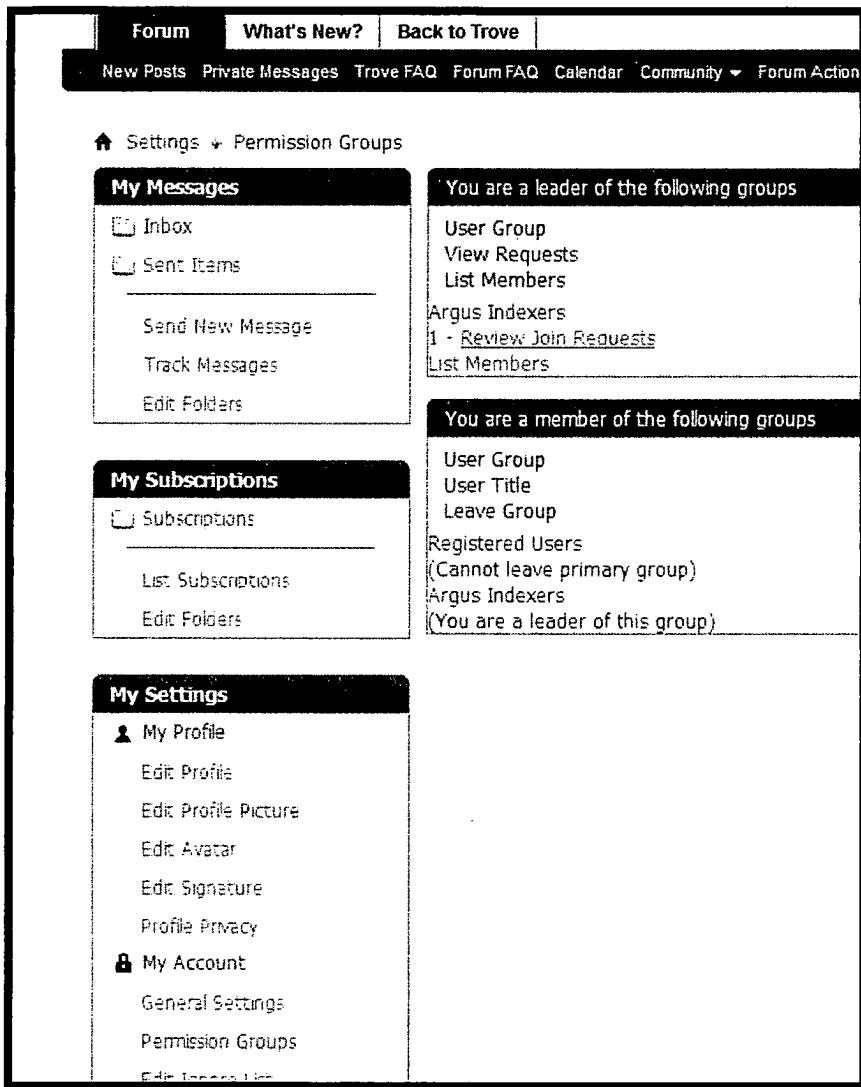


Figure 6. To approve a request to join a private forum, the group leader uses the "Settings > Permission Groups > Review Join Requests" menu.

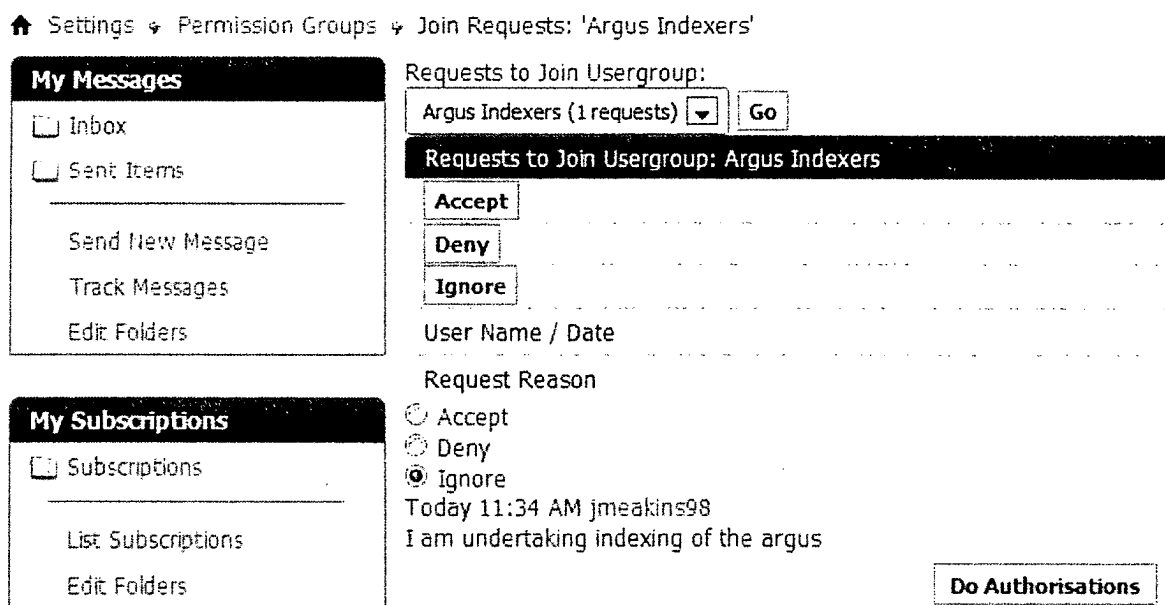


Figure 7. The group leader approves, ignores or denies requests to join a private forum

Method 3 – A password protected forum

Benefits:

- Argus Index owner can manage membership by emailing the password, which is very simple. There is no need to request of approve membership through vBulletin.

Drawbacks:

- NLA Div 3 staff need to do a small amount of set up.
- To block users, the shared password would need to be changed by NLA Div 3 staff.

NLA Div 3 staff would set up a new forum that requires a password to view. The forum will be listed with the other Trove forums (See Figure 8), but users will have to enter a password before they can see the forum contents (See Figure 9).

Anyone who knows the password will be able to access the forum. If the password needs to be changed, NLA staff will have to do this.

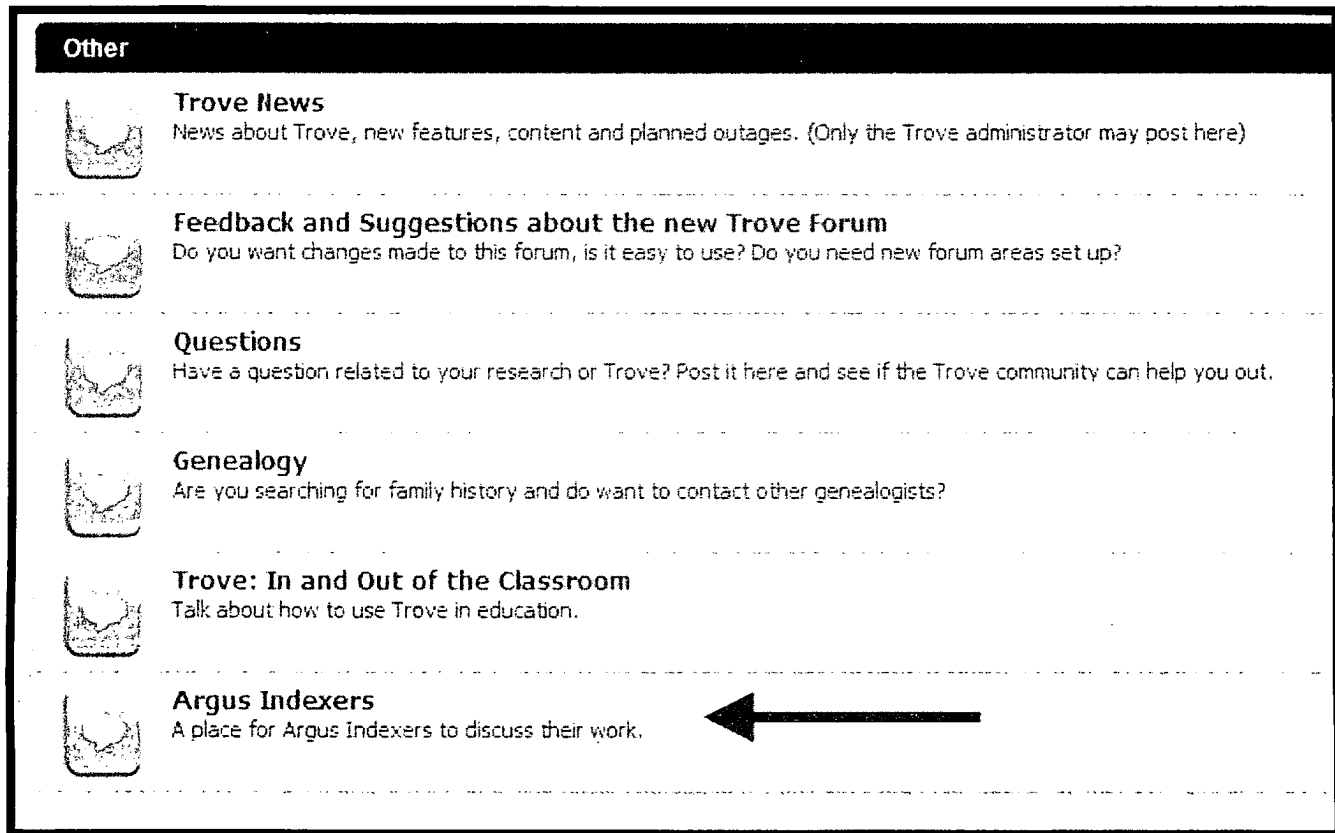


Figure 8. A password protected forum is listed with the other, open forums

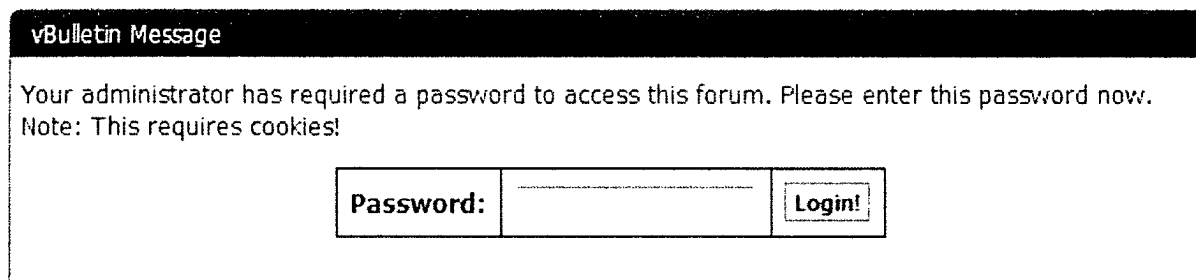


Figure 9. When users try to view the forum, they are asked to enter a password

Options for the future of the Argus Index (take 2)

Last Edited: 2 February 2012 by J. Meakins

Table of Contents

- Summary and guess of required time 2
- Option A. Add index terms to the digitised article, an interface to manage the terms, and a browse interface (was option 2) 3
 - Overview 3
 - Benefits 3
 - Drawbacks 3
 - Possible issues 3
 - Detailed description and mock-ups 3
 - 1. View index terms on an article 4
 - 2. Load index terms 5
 - 3. Edit index terms 6
 - 4. User management 9
 - 5. Browse 11
- Option B. Publish the index as a PDF, searchable in Trove 14
 - Overview 14
 - Tasks 14
 - Benefits 15
 - Drawbacks 15
 - Possible issues 15
 - Detailed description and mock-ups 16
- Option C. Retain the existing stand-alone application 19
 - Overview 19
 - Tasks 19
 - Benefits 19
 - Drawbacks 19
 - Possible issues 19
 - Detailed description and mockups 20

Summary and guess of required time

Option	Benefits	Challenges/Drawbacks
<p>A. Add index terms to the digitised article, + an interface to manage the terms + a browse interface</p>	<ul style="list-style-type: none"> - Preserves the structure and ability to browse the index - Index terms are added to the article, making them more discoverable - No ongoing data load - May facilitate inclusion of other indexes in Trove 	<ul style="list-style-type: none"> - Linking of index references to the correct article may not be accurate enough - Unclear whether this will fully cater to other indexes (further investigation needed) - Significant amount of work - Browse of index terms does not include the ability to search
<p>Guess: 19-20 weeks developer; 8 weeks interface</p>		
<p>B. Publish the index as a PDF, searchable in Trove</p>	<ul style="list-style-type: none"> - Preserves the structure and ability to browse the index - The index is searchable in Trove - If data is supplied as a PDF in future, ongoing data load is easy 	<ul style="list-style-type: none"> - Loading a PDF to DCM is possible but somewhat inelegant and unknown (but the PDF could be hosted elsewhere.) - Index terms are not located with the digitised articles - Div 1 are stakeholders but are yet to be consulted - Will indexers be able to deliver a PDF in future, including links? - Some links will lead to the wrong page, due to differences in the Argus version indexed and digitised
<p>Guess: 4.5 weeks developer</p>		
<p>C. Retain the existing stand-alone application</p>	<ul style="list-style-type: none"> - Preserves the structure and ability to browse the index 	<ul style="list-style-type: none"> - The index is not searchable or visible in Trove and so the value is lost to Trove users. - Users have to search Trove for the digitised article - Index terms are not located with the digitised articles - A separate application must continue to be maintained - Data must continue to be converted and loaded
<p>Guess: 5.5 weeks developer</p>		

Guesses of the time required to deliver a feature assume that a senior developer will be assigned. Further business analysis is required, whichever option is chosen.

Option A. Add index terms to the digitised article, an interface to manage the terms, and a browse interface (was option 2)

Overview

Index terms for the 1870s and 1880s would be automatically added to the digitised article in Trove, preserving the structure of the subheadings and navigation to other articles with the same index. The full list of index terms could be browsed.

New index terms for other years will be added and edited directly on the Trove interface, by users with special permissions. Some users may possibly be able to manage other user accounts.

This solution depends on being able to automatically identify the correct article.

Benefits

1. The structure and ability to browse the index is preserved
2. Index terms are added to the article, adding valuable keywords and making articles more discoverable
3. This solution may be able to be used by other indexing projects, and may inspire new indexing projects
4. Index terms can be added directly to the articles in Trove, there will be no ongoing data load for this index

Drawbacks

1. This option will take a significant amount of time to implement

Possible issues

1. Linking to the digitised article will not be completely accurate. It is still unknown whether an acceptable level of accuracy can be met.
2. The Trove interface will need to explain to users that only the Argus Index is included at the moment, and that not all years have been indexed.
3. If other indexes will be incorporated into Trove in future, investigation into the structure and requirements of those indexes is needed.
4. This solution has not yet been discussed with the Argus Index owner and may not suit her requirements, or may not be used in the way we expect (index terms added continuously, by many users).
5. There may need to be a way to search all index headings and to search sub-headings of a particular index heading. This functionality is currently available through the existing Argus Index Online.

Detailed description and mock-ups

The following sections are included below:

1. View index terms on an article
2. Load index terms
3. Edit index terms
4. User management
5. Browse

1. View index terms on an article

Index terms would be loaded to the digitised article, preserving the structure of the subheadings. The index terms will be similar to tags but will be more controlled and structured. Index terms would not be included in the tag cloud. They will not be editable by normal users. See Figure 1 for an example of how index terms may appear.

Navigation will be provided to the next and previous article with the index term (in chronological order). Clicking on an index term will show a list of all articles with that index term.

The screenshot shows the Trove website interface. At the top, it says "Trove Digitised newspapers and more" with navigation links for Home, About, Browse, and Help. Below this, the article title is "The Argus (Melbourne, Vic. : 1848 - 1956) about. Monday 29 August 1896". The main content area is divided into a sidebar on the left and the article text on the right. The sidebar contains several sections: "Lists (None yet)", "Tagged (None yet)", "Index Terms" (with sub-items: Warrnambool, laying of foundation stone at new masonic hall, Weather, heavy gale at Sydney), "Comments (None yet)", and "Electronically Translated Text". Arrows from text labels on the left point to these sections: "Click to view attribution and other info" points to the Lists section; "Click to view a list of other articles with this index term" points to the Index Terms section; "Click to view the next article with this index term" points to the Electronically Translated Text section. The article text on the right includes "TELEGRAPHIC DESPATCHES" and "QUEENSCLIFF, SATURDAY. The Pharos has towed the stern port quarter, and 50 feet of the deck, of the wrecked barque Sumex into Swan Bay." and "BELFAST, SATURDAY. Mr. P. J. Martin addressed a well-attended meeting of the electors of this division of the Western Province. He stated he was in favour of secular education in thinly-populated districts where one common school would meet the requirements of the inhabitants, but was in favour of the denominational system in the centres of population, where there were sufficient children to attend them. He was in favour of immigration; a five years' tenure of office for the members of the Upper House; the pink line of railway and harbour improvements. He was opposed to payment of members. He would like to see commercial men in the Upper House, for, had such been the case, no dead-lock would have ever occurred over the Darling grant. The candidate was well received. The schooner Mary Stewart left the wharf this morning, there being 8ft. 6in. on the bar, making the trip in and out in perfect safety."

Figure 1. An article displaying index terms (in a similar way to tags).

Tasks

Initially

- Write an information and attribution page for the Argus Index describing scope and coverage (linked from both the Argus title page and articles with index terms) (Business area is responsible)
- Implement attribution page
- Update the Trove API to include index terms with article metadata
- On the existing "View article" screen:
 - o Interface design to display the index terms
 - o Display the index terms
 - o Link index terms to a results page listing all the articles with that index term
 - o Link to the chronologically next and previous articles with the same term, in the same title

Guess of required time: 1.5 weeks developer: 3 days interface designer

When further data is received for the Argus Index

- No further data will be accepted. Indexers must add terms directly to Trove.

In future

- If other indexes are added in future, we would need to consider whether this interface meets their needs.

2. Load index terms

In order to show index terms on the digitised article, article references must be automatically matched to the digitised article. It is still unknown whether an acceptable level of accuracy can be met. There are several challenges which need to be considered.

- The Argus Index refers to the Melbourne edition of the Argus; however the NSW edition has been digitised in Trove. The two editions are not exactly the same, so the references in the Argus Index may not match up to the digitised articles.
- The Argus Index references the page number and column, which could contain more than one article.
- Not all words in the index term are included in the article text, and not all words in the digitised article have been correctly OCR'd.

Automatic matching for the index heading and reference:

abattoirs > increased water supply at city abattoirs 9 Feb 1870 5a [deputation]

might perform a search such as:

TITLE: The Argus (Melbourne) AND

DATES: 09/02/1870 AND

(abattoirs OR increased OR water OR supply OR city)

NOT advertising¹

although the exact query is yet to be developed. The matching software would then examine the results and find the article with the highest number of matching words, and the closest page number (and column) to that listed in the reference.

To manage problems with inaccuracies, a data load report should be produced automatically for the Argus Index owner (██████████) to review and action. The report should include:

- o terms which could not be matched to an article
- o terms which were matched to an article with a degree of uncertainty

Tasks

Initially

- Develop an article matching algorithm
- Load the existing Argus index data (1870s) and the new data on offer (1880s) into Trove, associating each reference with the correct article, in a way that also enables navigation, editing and browsing.²
 - o Quality assurance of data
 - o Design and implement database structures
 - o Modify newspapers lucene index
- Produce a report of non-matches and uncertain matches
- Update the Trove relevance ranking algorithm to incorporate index terms

Guess of required time: 8 weeks developer; 0.5 weeks BA (quality assurance)

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future, this data loading program may not necessarily be able to be reused.

¹

http://trove.nla.gov.au/newspaper/result?q=&anyWords=increased+water+supply+city+abattoirs&fromdd=09&frommm=02&fromyyyy=1870&todd=09&tommm=02&toyyyy=1870&l-title=|13&l-category=Article|category%3AArticle&l-category=Detailed+lists%2C+results%2C+guides|category%3ADetailed+lists%2C+results%2C+guides&l-category=Family+Notices|category%3AFamily+Notices&l-category=Literature|category%3ALiterature&l-word=*ignore*|*ignore*

² If a controlled vocabulary will be required in future, this solution must be built in a way that supports that functionality.

3. Edit index terms

An interface is needed to manage index terms. The interface should provide the following functionality:

- Add a new term
 - in a particular position
- Delete a term
- Change the indent level of a term
- Edit the wording of a term

Only authorised users may access this interface – not normal users. These users are divided into groups who each add and edit their own set of terms. For example, a group of indexers led by ██████████ may work on the Argus newspaper, using their own controlled vocabulary. A second group of indexers led by ██████████ may also be working on the Argus newspaper, but using a different controlled vocabulary. ██████████ indexers must be able to add and edit index terms created by any other indexers in their own group, but can not change index terms added by ██████████ group. (NOTE : This was a late breaking requirement and the rest of this document needs to be updated to take this into account)

The interface could be built into the existing article interface, or as a separate screen. Further interface design is still required. See Figure 2 for an idea of how this interface might appear to Trove users with appropriate permissions.

If the index terms needed to be limited to a controlled vocabulary, further complexity would be added. This has not been considered in the scoping of this task. It is considered briefly below under

The screenshot shows the Trove website interface. At the top, the Trove logo is displayed with the tagline "Digitised newspapers and more". Navigation links for Home, About, Browse, and Help are provided. A search bar is visible on the right. The main header indicates the newspaper is "The Argus (Melbourne, Vic. : 1848 - 1956)" and the date is "Monday 29 August 1870".

The left sidebar contains several sections for managing index terms:

- Lists (None yet)**: Includes a "Login to create lists" button.
- Tagged (None yet)**: Includes an "Add Tag" button.
- Index Terms**: Lists terms such as "Warrnambool" and "laying of foundation stone at new masonic hall." with a "Weather" section below it mentioning a "heavy gale at Sydney".
- Comments (None yet)**: Includes an "Add New Comment" button.
- Electronically Translated Text**: A section with a "Restricted" label, asking "Why may this text have mistakes?" and "How to correct this text?". It notes "No corrections yet" and includes a text input field.

The main content area displays the newspaper text, starting with "TELEGRAPHIC DESPATCHES. (FROM OUR OWN CORRESPONDENTS.)". The first section is dated "QUEENSCLIFF, SATURDAY." and describes a shipwreck: "The Pharos has towed the stern port quarter, and 50 feet of the deck, of the wrecked barque Sumax into Swan Bay." The second section is dated "BELFAST, SATURDAY." and describes a meeting: "Mr. P. J. Martin addressed a well-attended meeting of the electors of this division of the Western Province. He stated he was in favour of secular education in thinly-populated districts where one common school would meet the requirements of the inhabitants, but was in favour of the denominational system in the centres of population, where there were sufficient children to attend them. He was in favour of immigration; a five years' tenure of office for the members of the Upper House; the pink line of railway and harbour improvements. He was opposed to payment of members. He would like to see commercial men in the Upper House, for, had such been the case, no dead-lock would have ever occurred over the Darling grant. The candidate was well received." The final section mentions: "The schooner Mary Stewart left the wharf this morning, there being 8ft. 6in. on the bar, making the trip in and out in perfect safety."

Figure 2. An interface is needed to manage index terms

Exclusions

1. Inclusion of a controlled vocabulary. If this was added in future, questions to consider include:
 - a. Would the vocabulary terms be a suggestion or a limit?
 - b. Can the vocabulary be added to / modified?
 - c. Would the vocabulary apply to a particular group of users only?
 - d. Would a change to a term in the vocabulary retrospectively change all occurrences of that term?
2. Bulk changes to index terms.
3. A way for senior users to review index terms added by normal users.

Tasks

Initially

- Interface design and user testing for an edit interface
- Implementation of the edit interface

Guess of required time: 4.5 weeks developer; 6 weeks interface designer

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future, we would need to consider whether this interface meets their needs.

4. User management

User management requirements have not yet been discussed with the Argus Index manager (██████████). The assumption has been made that index projects will take ownership of a particular title (or possibly a set of titles if the periodical has ceased and continued under a different name). People outside the project must not be able to add or change index terms on those titles, as the projects use a very controlled vocabulary of index terms.

Two possible user management models are described below. It is not yet known which of these models is most appropriate for the Argus Index project, or whether these models will be appropriate for indexes incorporated into Trove in the future. It is possible that the Argus Index owner will not want other users to add index headings directly to Trove.

Model 1 – Flat structure, administered by NLA IT staff

Indexers

- Can add, edit and delete any index terms on a single title
- Are granted this role by NLA IT CAB staff

This model might be suitable for the Argus Index as only one title exists in Trove (not multiple titles). ██████████ currently reviews all indexing activity, so may prefer to be the only user adding the index terms to Trove, while the volunteer indexers continue to work in a paper based format.

This model is only suitable for a small group of users or a group of users that does not change regularly; otherwise it would require too much ongoing work for NLA IT CAB staff.

Tasks

Initially

- Require user permissions to access the index term edit interface
- Design and storage of permissions in the user management system

Guess of required time: <1 week developer

When further data is received for the Argus Index

- n/a

In future

- Add/remove indexer permission from users on request
- If other indexes are added in future, we would need to consider whether this model meets their needs.

Model 2 – Two tier, administered by index managers.

Index managers

- Manage a set of titles. For these titles, they can:
 - Add, edit and delete any index terms
 - Make other users *indexers* (and take this permission away)
 - Make other users *index managers* (and take this permission away)

Indexers

- Are assigned to the same set of titles as their index manager.
- Can add, edit and delete any index terms (on the titles they are assigned to)

An interface is required for index managers to give permissions to other users. This could be part of Trove, a standalone interface, or part of the pending Trove admin module. Figure 3 shows an example interface.

This model is suitable for a medium to large group of users, as it allows users to manage themselves.

Indexer user management

Titles

A list of titles for which you and these users can edit index terms.

- The Argus (Melbourne, Vic. : 1848 - 1956)

Add

Username: _____

e.g. johnsmith

Level: _____

Indexer

Add

Remove

Username: _____

avidtrover (indexer)

Remove

Figure 3. User management interface for index managers

Tasks

Initially

- Require user permissions to access the index term edit interface
- Design and storage of permissions in the user management system
- Interface design of user management interface
- Implementation of user management interface

Guess of required time: 1.5 weeks developer: 1 week interface designer

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future, we would need to consider whether this model meets their needs.

5. Browse

The Argus Index Online is primarily a browse-based interface. This allows users to get an impression of the events, news and topics reported on, in context. A browse interface for index terms could be added to Trove by listing index terms on the browse menu for the newspapers zone (see Figure 4).

The top index terms would be listed³, as well as an option to view all terms. If other indexes are added in future, this list could include terms from all indexes, not just the Argus Index.

There would be no change to the Argus title page in Trove. <http://trove.nla.gov.au/ndp/del/title/13>

The screenshot shows the Trove website interface. At the top, there is a navigation bar with 'Trove Digitised newspapers and more' and links for 'Home', 'About', 'Browse', and 'Help'. There are also 'Sign up' and 'Login' buttons. The main content area is divided into two columns: 'Browse Issues' and 'Browse Articles'.

Browse Issues

- by Title**
 1. The Argus
 2. The Mercury
 3. The Courier-Mail
 4. The Perth Gazette and Western Australian Journal
 5. The South Australian Advertiser
 6. The Australian Women's Weekly or Browse by cover (New!)

Show all titles
- by State**
 - National
 - ACT
 - NSW
 - VIC
 - QLD
 - SA
 - WA
 - NT
 - TAS
- by Date**

1803

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

Browse Articles

- by Category**
 - Advertising
 - Detailed lists, results, guides
 - Family Notices
 - News
- by Tag**
 - LRRSA (11221)
 - poem (5820)
 - BMD (4707)
 - Cricket (4081)
 - RWM (3681)
 - Cemetery of interment unknown (2745)
 - Rookwood Necropolis - interments (2608)
 - Cane (2534)
 - China (2511)
 - animation (2235)

Show all tags
- by Index heading** ←
 - Litigation
 - Labour
 - Accidents
 - Health - central board of
 - Ballarat
 - Land and works, department of
 - Gold
 - Meat

Show all index headings

Figure 4. Index headings could be listed on the browse menu for the newspapers zone

³ This may be difficult to generate if multiple titles are indexed

If users choose to view all terms, the top-level terms will be shown in an alphabetical list. Figure 5 shows a display similar to the “Browse headings” screen⁴ in the current Argus Index Online, including a navigation wheel and paging.

In future, if other indexes are added to Trove, a facet could be used to limit browsing of index terms to a single title only. Users may be required to choose a title first, due to implementation constraints.

Trove: Home About Site n

Trove Digitised newspapers and more
Home About Browse Help It's free and it only takes a minute [Sign up](#) [Login](#)

Browse index headings

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

▼ Title
The Argus (Melbourne ...

↑
Title facet could be added if other indexes are included

Aarons, J. 1872, 1 item
Aarons, Joseph 1870 - 1879, 49 items
Aarons, Lewis 1871, 3 items
Aarons, Solomon (money lender) 1870 - 1879, 9 items
Abattoirs 1870 - 1879, 189 items
 See also Echucha Slaughterhouse Company
Abbotsford 1870 - 1879, 10 items
Abbotsford Football Club 1878, 1 item
Abbott, David 1875, 1 item
Abbott, Henry 1872, 1 item
Abbott, H.H. 1878, 1 item
Abbott, Joseph Henry, 1877 - 1879, 2 items
Abduction, 1870 - 1879, 63 items
Abe, F.W. (Rev) 1873, 2 items
A'Beckett, - (Mrs) 1875, 2 items
A'Beckett, E. F. 1876-1879, 3 items
A'Beckett, L. J. 1870-1879, 19 items
A'Beckett, R. 1873, 1 item
A'Beckett, Thomas T. 1870-1879, 35 items
A'Beckett, W.A.C., 1871-1877, 13 items
A'Beckett, W. G. 1876, 1 item
Abel, Frederick Augustus (Prof.) 1871-1877, 2 items
Aberfeldie 1872-1876, 2 items
Aberfeldie Coursing Association 1878, 1 item
Aberfeldy River 1871, 1 item
Aberline, Joseph 1874, 1 item
Abernethy, J. M. (Rev.) 1870-1879, 10 items
Abington, Fred 1877, 1 item
Aborigines 1870-1879, 324 items
 see also Boomerangs, Church of England Mission to the Aborigines,
 Presbyterian Chinese and Aboriginal Missions

Next >

Figure 5. Index terms can be browsed

⁴ [https://www.trove.nla.gov.au/argus/index-headings](#)

When users click on a particular heading in the list, subheadings and references to articles will be shown (see Figure 6).

- Each article reference would be linked to the digitised article.
- A facet could be used to limit to a particular year (or decade) of references.
- Users could navigate back to the full list of top level index headings.

Again, if other indexes are added to Trove in future, a facet could be used to limit to index terms from a single title only.

The screenshot shows the Trove website interface. At the top, there is a navigation bar with the Trove logo and the text 'Digitised newspapers and more'. Below this, there are links for 'Home', 'About', 'Browse', and 'Help', along with a search bar and 'Sign up' and 'Login' buttons. The main content area is titled 'Index heading: Abattoirs'. On the left side, there are two facets: 'Title' (with a dropdown arrow and the text 'The Argus (Melbourne...') and 'Year' (with a dropdown arrow and a list of years from 1870 to 1879). An arrow points to the 'Title' facet with the text 'Title facet could be added if other indexes are included'. The main content area lists various newspaper articles related to abattoirs, such as 'Henry Walker, boiling down works nuisance 6 Jan 1870 5b [court case]', 'relocation of abattoirs 21 Jan 1870 supp 7g', 'overflow at Richmond abattoirs after storm 25 Jan 1870 supp 1d [meeting]', etc.

Figure 6. Browsing the "Abattoirs" index heading

Tasks

Initially

- Interface design of multiple browse screens
- Implementation of multiple browse screens
 - o Generate list of most frequent terms
 - o generate list of terms
 - o navigation
 - o decade and year facets
 - o link article references to the digitised article

Guess of required time: 4.5 weeks developer: 1.5 weeks interface designer

When further data is received for the Argus Index

- n/a

In future

- If other indexes are added in future:
 - o add a title facet
 - o add newspaper title name to the article reference
 - o consider whether this model meets the needs of the other indexes

Option B. Publish the index as a PDF, searchable in Trove

Overview

The 1860s decade of Argus index data was published as a book⁵. We could migrate each year of index data currently in the Argus Index Online into an online PDF. The PDF could be hosted by the NLA with the full text searchable in Trove (similar to finding aids and “local history resources” indexed by Trove) and index references linked to the digitised page in Trove.

The layout of the PDF could be styled on the layout of the existing book.

The existing decade (1870s) and the next decade of data on offer (1880s) could be converted to PDF by the NLA. In future, the NLA would only accept the data in PDF format. Digitisation of the 1860s hard copy index is not considered here.

NLA cataloguers would need to create a metadata record in the NLA catalogue for each PDF – one for each year included. (The metadata records will automatically flow through to Libraries Australia and Trove.)

Tasks

Initially

1. Design the layout of an Argus Index PDF (responsibility unclear)
2. Convert the existing Argus index data (1870s) and the new data on offer (1880s) into text searchable PDFs – one for each year, with article references linked to the digitised page in Trove.
3. Upload the Argus Index PDFs to DCM (or another public location).
4. Create metadata records for each Argus Index PDF in the NLA catalogue, including a link to the PDF and a flag to indicate that the PDF should be indexed by Trove. (Cataloguing staff are responsible, possibly some consultation with IT)
5. Organise for data to be provided in PDF format in future. (Business area is responsible)

Guess of required time: 4.5 weeks developer

When further data is received for the Argus Index

1. Upload the Argus Index PDFs to DCM (or another public location).
2. Create metadata records for each Argus Index PDF in the NLA catalogue, including a link to the PDF and a flag to indicate that the PDF should be indexed by Trove

Guess of required time: <1 day developer

In future

1. An experimental project could be carried out to investigate adding index data as tags on articles. This project could examine the feasibility and usefulness of this approach. It is recommended that this project:
 - a. uses an index that applies to the same edition of the newspaper as that digitised in Trove, to facilitating matching index terms to the correct digitised article
 - b. involves an index owner who is amenable to converting the index to tags
2. The hard copy Argus index from the 1860s and other years could be digitised, OCR'd and made searchable in Trove. References to these indexes can be found here:
<http://trove.nla.gov.au/book/result?q=%22argus+index%22>

⁵ <http://trove.nla.gov.au/work/32646197>

Benefits

1. The structure of the index and the ability to browse it is preserved.
2. This solution does not rely on being able to automatically link an index reference with the corresponding digitised article in Trove.
3. The index is searchable in Trove.
4. If data is supplied as a PDF in future, ongoing data load is easy.

Drawbacks

1. The Argus index terms are not added to the digitised article, so do not add additional search terms for users.
2. The Argus Index is searchable in a different zone to the Argus newspaper articles.
3. Some of the page references will link to a page which does not show the article – due to differences in the Argus version indexed and digitised.

Possible issues

1. This solution assumes that a single year of the index is not too large to cause problems in DCM, for the Trove indexer, or for users to open. (If problems arise, the PDFs could be more granular.)
2. The only capability DCM has for uploading PDFs is as a finding aid (PDFs of books are automatically generated from images and do not include text). The Argus Index may need to be uploaded as a kind of finding aid, including a finding aid PI. DCM was modified to handle PDF finding aids in 2011 but this functionality has not yet been used. Alternately, the PDF could be hosted outside of DCM. Also note that the NLA's DLIR project is due to deliver a solution to upload and deliver PDFs with OCR'd text by end 2012.
3. Will indexers be able to deliver a PDF in future, including links to the digitised page in Trove?
4. NLA Div 1 are stakeholders but have not yet been consulted.

Detailed description and mock-ups

Users can search Trove for words included in the index, but Trove will only indicate whether their search terms are found in a particular year of the index. The results will appear in the books zone. Users click on the “view online” link to view the PDF of the index. (See Figure 7)

NATIONAL LIBRARY OF AUSTRALIA

Home

Trove

All **Books** Pictures, photos, objects Journals, articles and data sets Digitised newspapers and more Music, sound and video Maps Diaries, letters, archives Archived websites (1996 – now) People and organisations

lightning accident kiandra **Search**

Available online Australian content In my libraries [Advanced search](#) [Search tips](#)

Refine your results:

Books

Showing: 1 - 20 of 8,805 Sort by: Relevance **Sort**

Format

- Book: 8,788
- Braille: 5
- Illustrated: 200
- Large print: 12
- Audio book: 17
- Conference Proceedings: 2
- Thesis: 1

Availability

- Online: 8,788
- Access conditions: 14
- Freely available: 6,600
- Unknown: 1

▶ Decade

▶ Language

▶ Australian Content

2. Various years of the Argus Index may appear in the list of results

3. Click to view the index

Proceedings relative to the accident by lightning at Heckingham
[Book: 1783]
At 5 libraries

Effects of lightning upon the human body, exemplified by a remarkable accident / by Dr. Tilesius
Tilesius von Telenau, Wilhelm Gottlieb, 1769-1857
[Book: 1824]
Keywords: Electrical injuries., Lightning
At Baillieu Library

Indexes to the Argus, 1860 / indexed by Geraldine Suter
Suter, Geraldine
[Book: 1999]
Keywords: Argus (Melbourne, Vic.) - indexes, Australian newspapers - indexes, Newspapers - indexes.
...four killed in lightning accident at Kiandra, 1860-09-20 pg5 col f men
fall down holes at Inglewood 1860-09-24 pg 5 col g...
 View online
At National Library

A Full and true relation of the death and slaughter of a man and his son at plough, together with four horses, in the parish of Cookham in the county of Berks. Sept. 2, 1680 : slain by the thunder and lightning that then and there happened, as may fully be testified by credible persons whose names are

Figure 7. Finding the Argus Index as a result in Trove

Users then search the PDF for the terms they are interested in (see Figure 8). Users can also browse the PDF of the index if they wish.

Accidents

see also Drowning; Fires; Industrial accidents; Mining accidents; Poisons; Railway accidents; Road accidents; Safety; Shipping accidents; Shooting accidents

1860 -04-07 pg 6 col c
driver from Mount Esse station killed in dray accident,
1860-03-02 pg 3 col g
glass globe falls on man in Legislative Assembly,
1860-03-16 pg 5 col b
strange accidents at Theatre Royal, 1860-03-17 pg 5 col c
child swallows seeds of thornapple at Ballarat, 1860-03-21 pg 5 col c

man loses eye in accident with bullock whip, 1860-03-31 pg 5 col g
old man, brought to court at Maryborough on assault charges, injured, 1860-04-04 pg 5 col c
child falls into waterhole at Malmsbury, 1860-04-17 pg 6 col f
boy falls down hole at quoits match, 1860-05-11 pg 5 col g
man injured in earthfall at Prahran railway works,
1860-05-16 pg 5 col a
man dies at Bendigo after falling under load of wood,
1860-06-04 pg 4 col d
woman dies from asphyxiation while drunk, 1860-06-13 pg 5 col g
Buninyong man rescued by Chinese three days after falling into hole 20 ft deep, 1860-06-16 pg 5 col d
part of cornice of new Bank of Victoria building at Hamilton falls to ground, 1860-06-19 pg 5 col c
man killed in blasting operations at Kyneton, 1869-06-29 pg 4 col b
sailor dies due to lack of lighting on Sandridge railway pier, 1860-08-20 pg 6 col e
child suffocates on piece of meat, 1860-08-21 pg 5 col b
child dies after falling down shaft at Inglewood, 1860-09-17 pg 5 col c

drunken man in accident at Theatre Royal, 1860-09-18 pg 5 col d
four killed in lightning accident at Kiandra, 1860-09-20 pg 5 col f
men fall down holes at Inglewood, 1860-09-24 pg 5 col g

Brown, John, Yackanda over stump, 600615 7
Burke, Patrick, killed in Burrigge, -, dies after f 600917 5c
Butt, George, attorney's springcart in Little C Campbell, -, mounted c horse at Morses Cree Caulivan, Michael, run Chaplin, George, falls f Christopher, James, wo powder, 601121 5a
Clarke, G., Maldon, ac 601106 5d
Coard, -, killed in hors Coates, -, dies after fal Cooke, William, kicked Cox, -, thrown from ca Davis, Henry, scalped i Creek, 600605 7a
Deuning, Alfred, killed 600702 4e
Donald, John, injured t 600221 5b
Dudden, Alfred Henry, 600818 5c
Earl, Sarah, falls off ca Easton, James, run ove Erskin, -, thrown from Falkner, George, sheph for family, 601022 5
Faris, Frank B., Emeu lark after hunt, 6006
Farrel, Edward, injured

Figure 8. Searching a PDF of the Argus Index

When a user finds a reference to an article they want to read, and click on the page link to return to Trove. The user must then find the article on the page (See Figure 9). In some cases, the article will be located on a different page.

The screenshot displays the Trove website interface. At the top, there is a navigation bar with links for 'Home', 'About', 'Browse', 'Help', 'It's free and it only takes a minute', 'Sign up', 'Login', 'kiandra', 'Search', 'Show search limits', and 'Advanced search'. Below this, the page information indicates 'The Argus (Melbourne, Vic. : 1848 - 1956) about' and 'Thursday 20 September 1860' on 'Page 5' of 8. A sidebar on the left lists 'Page 5 articles' with four entries: 'THURSDAY, SEPTEMBER 20, 1860.' (4960 words), 'TELEGRAPHIC DESPACHES. (FROM OUR OWN CORRESPONDENTS.) SYDNEY, WEDNESDAY.' (163 words), 'SANDRIDGE DOCKS AND TRAMWAY COMPANY. TO THE EDITOR OF THE ARGUS.' (231 words), and 'AGRICULTURAL REPORT. WEDNESDAY, SEPTEMBER 19.' (The produce markets are improving again). The main content area shows a preview of the 'AGRICULTURAL REPORT' article, which discusses the state of the produce markets and the weather. The article is titled 'AGRICULTURAL REPORT.' and begins with 'The produce markets are improving again...'.

Figure 9. The Argus Index could link to the digitised page in Trove. Users must then find the article.

Option C. Retain the existing stand-alone application

Overview

As the current server is being decommissioned, the existing application could be ported to the new server and infrastructure.

The decade of data currently on offer would be loaded into the existing application. A decision would need to be made as to whether the National Library continues to host the application indefinitely, loading data periodically as it is made available, or not.

Tasks

Initially

1. Port the existing application to the new server and infrastructure
2. Load the new data on offer (1880s) into the existing Argus Index application
 - a. Document how to load the data (in svn)
3. On the Argus index Online home page, add a link to the Argus title page in Trove

Guess of required time: 5.5 weeks developer

When further data is received for the Argus Index

If the NLA decides to continue loading new data:

1. Load the data into the existing Argus Index application

Guess of required time: 1.5 weeks developer

In future

If the NLA decides to continue to host the application:

1. Continue to maintain the Argus Index application
2. Consider enhancing the Argus Index to link article references to the digitised page in Trove

Otherwise:

1. Decide what to do with the Argus Index (again)?

Benefits

1. Preserves the structure of the index
2. Preserves the ability to browse the index

Drawbacks

1. The index is not searchable or visible in Trove
2. Users have to search Trove for the digitised article
3. Index terms are not located with the digitised articles
4. A separate application must continue to be maintained
5. Data must continue to be converted and loaded

Possible issues

1. The existing Argus Index application has no current business owner

Detailed description and mockups

The existing Argus Index application will continue to exist (see Figure 10). It includes the ability to:

- search the index terms
- browse the index terms alphabetically
- browse the index terms alphabetically for a particular year

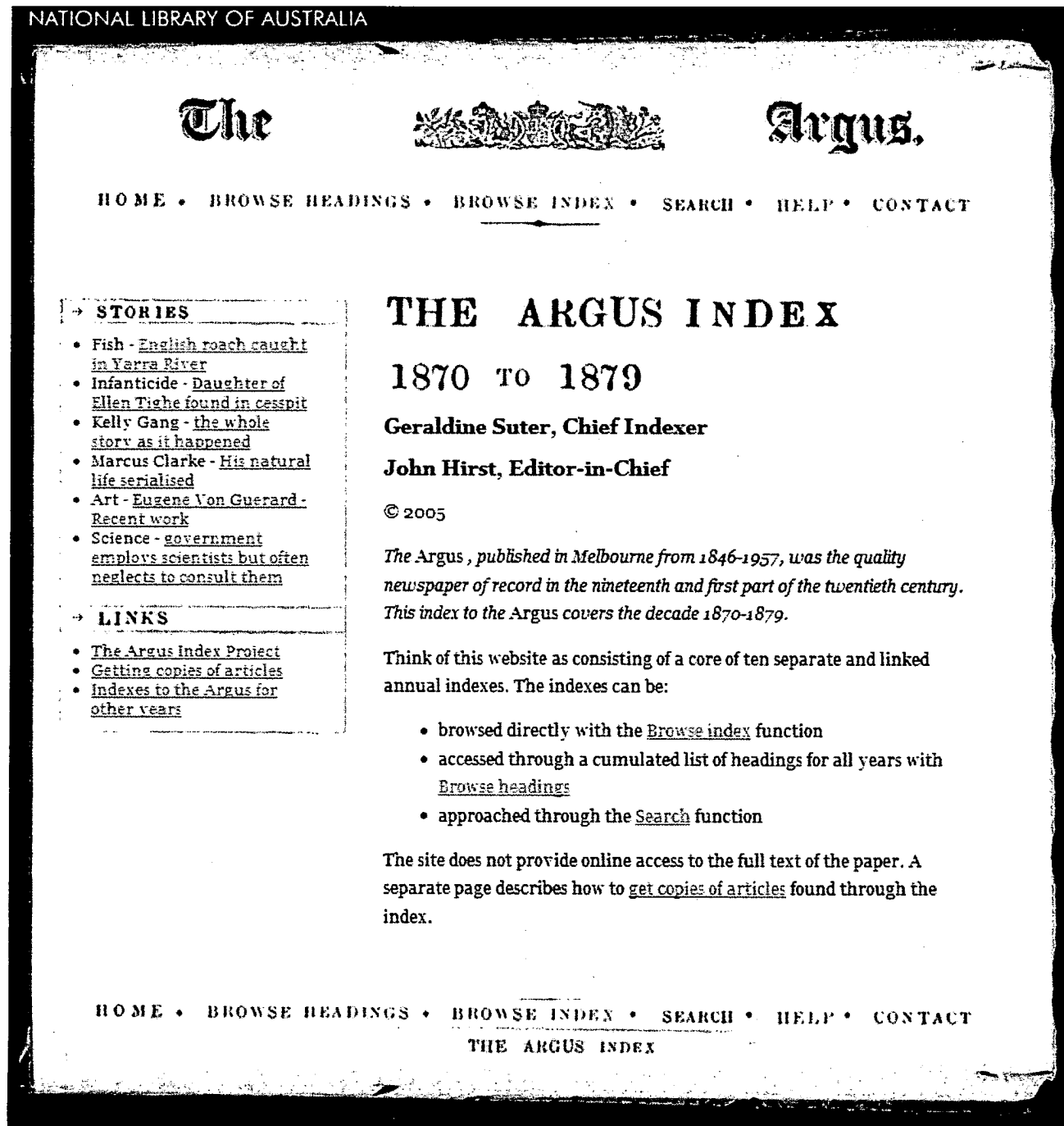


Figure 10. The existing Argus Index application

[REDACTED]

Subject: Making a final decision on the options for the Argus Index Online
Location: Libraries Australia Meeting Room

Start: Thu 2/02/2012 11:15 AM
End: Thu 2/02/2012 12:15 PM

Recurrence: (none)

Meeting Status: Accepted

Organizer: [REDACTED]

Required Attendees: [REDACTED]

When: Thursday, 2 February 2012 11:15 AM-12:15 PM (UTC+10:00) Canberra, Melbourne, Sydney.
Where: Libraries Australia Meeting Room

Note: The GMT offset above does not reflect daylight saving time adjustments.

~~*~*~*~*~*~*~*~*

In light of the discussion at yesterday's Trove Reference Group meeting, it has been decided to further examine the options for the future of the Argus Online Index.

The following reading is recommended prior to the meeting:

- The paper presented to the TRG is at [REDACTED]
- The summary of the most recent inter-Divisional discussion at [REDACTED]
- The options paper prepared by Joanna Meakins with input from the two Trove teams at [REDACTED]

Thanks,
[REDACTED]