

16 January 2012
Council Room

TRIMmed
DATE:
No: R12/31446
p7

Present:

Apologies:

Briefing papers:

- Item 2: Trove Governance – CMG decisions
- Item 5: Proposed Terms of Reference for Trove Reference Group
- Item 6: Proposed meeting schedule
- Item 7: Proposed internal communication strategy
- Item 8: Trove scope statement – draft for consideration
- Item 9b: Inclusion of newspaper indexes

Outcomes/ **2. Trove Governance – CMG decisions**

Discussion

confirmed that the recommendations made in the Trove Governance paper had been accepted by CMG. The key points to note were:

- The formation of the Trove Reference Group (TRG), with broad corporate representation, was confirmed by CMG;
- The Trove business owner is the Director, Collaborative Services Branch;
- Divisional representation is critical to the success of the Group; and
- The TRG will take its recommendations to CMG for final decisionmaking.

It was agreed that responsibilities will be revisited in 12 months time.

3. Trove teams in Divisions 3 and 4.

and introduced the Trove teams in Divisions 3 and 4.

The Trove Team for the Resource Sharing Division consists of:

- as the Trove business owner;
- A Trove Manager, currently ;
- , who is responsible for biographical data contributors and oversight of the decommissioning of Music Australia;
- , who works with other data contributors and provides enquiry responses;
- who provides enquiry responses, particularly relating to the Newspapers zone, manages fora and other social media interactions;
- is working on a productivity project to identify a solution for a new Trove Vendors and Contributors database to manage organisational contributors to Trove; and

The Libraries Australia Customer Services team also often provides a first point of contact and promotional services.

Division 4 has no single Trove team. Responsibilities for support of different aspects of Trove fall to different teams, as follows:

- Collection Access Branch, managed by : code base and support for Trove data management issues;

- IT Services, managed by [REDACTED], is responsible for infrastructure and day-to-day operational support of the service, capacity planning and related tasks;
- Collection Infrastructure Branch, managed by [REDACTED], is responsible for Digital Library and Library operations systems, e.g. the Newspapers system which is closely aligned with Trove. With Digital Library Infrastructure Replacement (DLIR) project it is anticipated there will be more interaction between Trove as a discovery and access mechanism for collection items;
- [REDACTED] reported that [REDACTED] is the main IT support person for the Newspapers Content Manager (NCM), and she is the main support BA.

4. Trove Business owner report.

a. Managing Contributors

[REDACTED] reported on the Trove Vendors and Contributors database project currently being managed by [REDACTED]. This is one of the Library's 2011/2012 positioning projects. A briefing note about this project is available at [REDACTED]
[REDACTED]
[REDACTED]

The project includes an investigation of systems already used in the library, and the ideal solution would integrate with other services and require only modest development. It was noted that the common problem of systems not integrating fully would probably not be resolved; however the driver for this project was that under the current system for management of contributors it is impossible to report on the status of contributors for prioritisation and consolidation of business information.

The IT Division sought further information about the project prior to the recommended decision being presented to CMG.

b.i. Application Programming Interface

[REDACTED] reported on the Trove Application Programming Interface development. The briefing note on the API is available from: [REDACTED]
[REDACTED]

[REDACTED] asked if individuals would have access to the API. [REDACTED] clarified that individual access to the API would be supported and that the first individual will be [REDACTED], who is scheduled to research data mining as a Harold White Fellow in early 2012. It was noted that there may be commercial issues associated with the use of the API, which will need to be resolved for paid content such as the Australian Joint Copying Project AJCP.

[REDACTED] noted that the Australian Women's Weekly pages will not be available to download via the API due to copyright reasons. [REDACTED] confirmed that each additional exception added to the API adds a burden when processing, and the preference would be that any in-copyright newspaper content to be included in Trove would also be fully available through the API. She also noted that a few people are helping themselves to the content anyway, so it is preferable for the content to be made available properly rather than adding additional rules.

[REDACTED] confirmed that the sharing of the AWW images was not approved by Australian Consolidated Press, which is why they are not included in the API. He also recommended that the Library negotiate with contributors to ensure that all material available through the Trove web interface be made available through the API. [REDACTED] advised that she will need to discuss the issue of post-1954 newspapers with the state and territory libraries as there had been recent negotiations for new content of this type.

[REDACTED]
[REDACTED]

b.ii. Decommissioning of Music Australia, Picture Australia and Australia Dancing

[REDACTED] provided a report on decommissioning of Music Australia, Picture Australia and Australia Dancing. A briefing note on the decommissioning of these services is available at:
[REDACTED]
[REDACTED]

[REDACTED] asked about development of discipline views for Trove, as this work had been postponed. The discipline views will be discussed on the Trove planning day. It was noted that expectations from the music and dance communities would need to be managed. Preparatory work for decommissioning could be shared with the TRG before the planning meeting.

5. Terms of Reference

[REDACTED] introduced the draft Terms of Reference for the TRG.

ACTIONS (Division 3):

- 1) It was agreed that the following amendments would be made:
 - Dot points under members referring to divisional representatives to be combined;
 - Add dot point to allow that staff members with a specific expertise may join the meetings for a specific agenda item;
 - Delete the Quorum requirement;
 - Ask for agenda items to come forward to secretariat at least 10 working days before a meeting, which will be followed by discussion on whether a paper is needed;
 - Agenda and papers will be sent out five working days prior to the meeting.
- 2) Division 5 will be asked to provide a representative to future meetings.

6. Proposed meeting schedule

In discussion of the proposed schedule of quarterly meetings including an annual planning day, it was agreed that a half day should be sufficient for the planning day. However this should be booked for a full day, only using the time required.

The timing of the planning meeting was also discussed, with [REDACTED] asking whether the outcomes of the planning day would flow through to the workplans for the following financial year. A 5 March planning meeting would occur 11 months before funding decisions are usually made. [REDACTED] clarified that CMG sees the TRG as a group dealing with medium-term priorities.

[REDACTED] suggested that it may be appropriate to have a special purpose meeting soon after the agreed March planning day to clarify the priorities and outcomes. This will be trialed in 2012, with a meeting to occur two weeks after the planning day. It was also agreed that for 2012 the other planned meetings would be kept as scheduled.

7. Proposed internal communication strategy

Meetings, the use of the shared J drive and the TRG email were confirmed as proposed. The trg@nla.gov.au email address can be used to raise issues with the Trove business owner.

It was also agreed that active communication between members could be managed via a distribution list, and that a wiki is not currently needed.

8. Trove scope statement

A draft scope statement presented to the TRG was based on the Libraries Australia scope statement [REDACTED]. A collection development policy will also be developed.

The draft statement did not include information about who Trove is for or the aspirations of Trove, and that it appeared to be a statement of content rather than a statement of scope. This is an important distinction as Trove is a service, and is much more than just the content. [REDACTED] also raised concerns that the language and description within the draft statement may not be appropriate for the intended audience of the statement, with many new providers of content to Trove not from a library background.

It was agreed that it would not be possible to capture all relevant points in a single document, and that it would be necessary to have multiple documents, one with a broader statement for everyone, and a second document aimed at libraries that would clarify the difference between Libraries Australia and Trove.

[REDACTED] mentioned that Division 5 had done work on simplifying information about Trove for a general audience; however the resulting document had raised unanswered questions regarding the difference between Trove and Google. [REDACTED] raised that caution needs to be used in statements about the size of Trove as a resource discovery tool, as Trove is indexed by Google, and in the current environment Google will remain the largest single source for resource discovery.

From a marketing perspective, the points about full-text content need to be made more prominent. It was also agreed that, with the addition of points about Oral History and Manuscripts, the current draft statement is an appropriate content overview rather than a scope statement.

[REDACTED] mentioned that the use of 'freely available' in the content statement may be misinterpreted by users, and that the term 'freely available' should not appear in the statement. [REDACTED] mentioned that there had been recent discussions in Division 1 about freely available resources, and how access to these would be provided by the National Library, and that the conclusion was that Trove and/or Libraries Australia would be the most appropriate paths for providing this access. This does need to be examined in more detail, however, from collection development, database, and funding perspectives. These are some of the aspects that need to be considered as part of the development of a Trove Collection Development Policy.

ACTIONS (Division 3):

- 1) The draft statement will act as a content overview;
- 2) A statement about Oral History transcripts and Manuscript finding aids will to be added;
- 3) The full text content that is available in Trove will be highlighted;
- 4) a second document describing the purpose, aspirations and audiences of Trove will be developed;
- 5) A Collection Development Policy, to address issues of where content should be disclosed, will be drafted.

[REDACTED] committed to having new draft documents available for consideration on the planning day, including an overview of the questions that need to be addressed in the Collection Development Policy.

9b. Inclusion of newspaper indexes in Trove

[REDACTED] notes on the inclusion of newspaper indexes in Trove, to be considered in addition to the paper, are available at [REDACTED]

The creators of the Argus Index provided a new set of index data to be loaded into the separate Argus Index website. As it has not been updated for some time, the impetus to load the data was discussed, and it was decided to take the opportunity to look for alternative options. Due to the length of time that the service has been available, Division 3 considers that simply turning off the [REDACTED]

edition has been indexed, so the index is essentially to an off-line work. This could be achieved by storing the index as a digital object that contained a record in Trove, but was not stored in Trove itself. In discussion, the disadvantage of this approach was that the vast majority of users access Trove's newspapers only, so the data would not be seen by these users, and Option 2 increases the value by adding the terms where newspapers users will benefit from them. It is noted that the cost of option 2 still needs to be assessed.

Other index services (e.g. Gale) sometimes have full text and are not loaded into the Journals zone, as with other journals. There is concern about the user experience with the separation of data, and confusion around what is in each zone. In discussion this was clarified to mean that users are having trouble finding what they want in some cases, as the information is contained in different sections of Trove. Users who enter through the Newspapers homepage do not see the content, and users who land on the main page have to remember which section the information they were seeking was contained in. The first point of contact for confused users was not normally the FAQ. The separation of data is due to content placement, with the newspapers zone containing the full text of the newspaper articles, whereas the other zones contain metadata only.

It was agreed that if there is useful data available, then the effort should be made to get this data into Trove, and that if there is an index for digitised content then users would expect there to be a link to the content in Trove.

ACTIONS:

- 1) Estimate cost and timeframes to develop Option 2;
- 2) Division 4 to document an alternative option;
- 3) [REDACTED] to contact [REDACTED] with options for the Argus Index.

10a. Forward Agenda for planning meeting on 5 March 2012

It was agreed that the following items would be on the agenda for discussion at the planning meeting:

- 1) Discipline views and their implications.
- 2) How to handle indexes to digitised full-text content.
- 3) How to handle indexes to other content.
- 3) A draft Scope statement.
- 4) Overview of the anticipated Collection Development Policy.
- 6) How to expose Oral history/audio in Trove, and encouraging other institutions to adopt NLA schemas and standards.
- 7) PANDORA decommissioning and how to provide access to archived webpages through Trove.
- 8) Presentation on the progress of decommissioning work for Music Australia, Picture Australia and Australia Dancing, including the replacement biographical tool.
- 9) Briefing on the Libraries Australia Search redevelopment.

Other items that will need to be discussed in future meetings include:

- 1) Assessment of the Journals, articles and data sets zone and its success after 12 months.
- 2) Re-examining the Journals, articles and data sets zone against outstanding recommendations and actions.
- 3) Assessment of Trove's role in respect to the DLIR, including whether the discipline view would include NLA digitised journal and book content and whether newspapers would be pulled back from Trove as a separate discovery service.

Meeting closed at 3.55pm.

[REDACTED]

[REDACTED]

MEMORANDUM

DATE 10 January 2012
REF NLA11/1865
TO Trove Reference Group
FROM [REDACTED]
SUBJECT **The potential inclusion of newspaper index data into Trove**

PURPOSE

To discuss, and if possible, confirm as part of Trove's scope, whether the addition of newspaper index data is a valuable strategic direction to take.

BACKGROUND

The National Library has hosted the Argus Index Online service since 2002. A team of volunteers is used to create the index data, which progresses at a modest pace. In mid 2011, the Library was offered four more years of data (1880-1883) to add to the decade of the 1870s already available in a separate interface at www.nla.gov.au/argus/.

The requirement to make a decision about the future of the Argus Index Online was brought to the Resource Sharing Division to determine whether it could be incorporated into Trove. In the Issues section below, there is a summary of the options considered.

If inclusion is not considered viable, then a decision about the future of this and other newspaper indexes will need to be made corporately. For example, in 2011 the National Library was offered a copy of the index to "Walkabout" magazine for inclusion in Trove.

A letter was also provided by the managers of the Argus Index Online data outlining their requirements (attached).

ISSUES

The code used to load the index data originally has not been utilised for several years, and does require some (unscheduled) work by the Information Technology Division to reactivate. Its application would need to be agreed directly with the data owners. This was taken into account during further discussions.

Following some work by the Information Technology Division to scope possible technical solutions for the inclusion of the Argus Index into Trove, a meeting was held in December 2011 to discuss the options proposed. The options are discussed in full at R11/113832ⁱ, but may be summarised as follows:

One – add a record to Trove for each major heading:

Argus Index - Warrnambool

Geraldine Suter, John Hirst

[Article: 1870-1879]

Keywords: Warrnambool

... at port 5 May 1871 p.5 [celebration] doctor or sea porcupine found 18 May 1871 supp p 1 woollen manu...

[View online](#)

Two – add index annotations (not tags or comments) to the digitised article:

The Argus (Melbourne, Vic. : 1848 - 1956) about Monday 29 Aug

Three – add index data to the digitised article as tags:

Four – create a specialised interface for viewing the index, powered by Trove:

Five – in addition to Option 1, add an extra zone in which to view index content:

Six – retain the existing stand-alone application. One other consideration discussed was that the existing user interface needs to be moved from its current server by 30 June.

The Resource Sharing Division advised that it would welcome the inclusion of the newspaper index data into Trove in line with Option 2. It was also noted that the Resource Sharing Division could not assume responsibility for the future of the index data if it was not integrated into Libraries Australia or Trove.

Note that inclusion in Trove may be subject to the resources required to achieve the enhancement, and prioritisation of all enhancement activities by the TRG at a future meeting.

RECOMMENDATIONS

1. That the Trove Reference Group confirm its interest in the inclusion of high quality controlled vocabulary indexing of full text material into Trove, confirmed by a clear cost-benefit articulation.
2. If recommendation #1 is agreed, that the Argus Index online data be included in Trove according to option 2, preferably by using a model which can be redeployed for other indexes.

ATTACHMENT

Letter from [REDACTED] and [REDACTED] outlining their future requirements for the Argus Index Online.

[REDACTED]
National Library of Australia
Canberra ACT 2600

6 December 2011

Dear [REDACTED]

We understand from our telephone conversation with [REDACTED] on 21 November that the NLA is soon to make decisions on the future of *Argus Index Online*. We offer this information on our index and our views on its future in Trove. We look forward to further discussions with you before any final decision is made.

Retention of existing service

The goal of the Argus Index Project is to produce the first complete index to a major Australian daily comprising born-digital indexes for the period 1870 to 1909 and the conversion of the existing hard copy indexes for the periods 1846 to 1859, 1860 to 1869 and 1910 to 1949 to the online format. We are very happy with the current online version which provides *enhanced access to the index while preserving the functionality and integrity of the printed indexes in which each year is a self-contained work*¹.

