

MEMORANDUM



DATE 5 March 2009
REF NLA07/1529
TO ANPLAN
FROM Rose Holley – Manager Newspaper Digitisation Program
SUBJECT **Progress Report on the Australian Newspapers Digitisation Program August 2008 – February 2009**

Purpose

This report summarises progress on the Australian Newspapers Digitisation Program (ANDP) over the last six months from August 2008 – February 2009 and outlines work to be completed in the next 6 months.

1. Work progress August 2008 – February 2009

- Evaluation of 2 Request For Tender (RFT) submissions for OCR and Related Services and Scanning Services was undertaken between August and November 2008. Approval to enter into a Deed of Standing Offer with the preferred Tenderers was provided by the Minister for Environment, Heritage and the Arts on 16 December 2008. Contract negotiations have been undertaken and the Library is now in the process of appointing a panel of contractors. When contracts are finalised the Library will publicly announce the panels.
- 450,000 newspaper pages were digitised (scanned) and quality assured in the period. The number of pages scanned to date is 1.85 million.
- 300,000 newspaper pages were OCR processed and delivered to the public in Australian Newspapers beta service in the period. The total number of pages available to the public is 367,651.
- The Australian Newspapers beta service was available to users throughout the period (and will continue to be so). Feedback and usage statistics were gathered from users over a 6 month period. The feedback and usage figures were collated and made publicly available on the ANDP website on 24 December 2008. See http://www.nla.gov.au/ndp/project_details/ Overall the service is being heavily used and a significant volume of positive feedback has been received. Around 1300 members of the public have actively corrected 100,000 articles (2 million lines of text) and added 45,000 tags to articles.

- The Australian Newspapers beta service was enhanced. A frame splitter function was added to enable users to have control over the viewing width of the newspaper page image and the OCR text.
- A review of titles selected was undertaken with stakeholders during the period with the aim of expanding the list. An extended title list was approved by all stakeholders and made available to the public in December. See:
www.nla.gov.au/ndp/selected_newspapers/documents/ANDP_Selected_Titles_Jan_2009.pdf
- State and Territory Libraries nominated a representative for liaising with organisations and individuals in relation to contribution of titles. The representative will also review public suggestions for titles for their states.
- In consultation with stakeholders a model for national contribution of additional newspaper content was developed. An Australian Newspaper Digitisation workshop was held in November 2008 for stakeholders where draft contributors guidelines were reviewed and discussed.
- A list of potential titles for contributions from each state/territory was collated. Stakeholders requested this information not be made public at this stage.
- Work began in January 2009 on modification of the Newspapers Content Management System to enable contribution of content by stakeholders.
- Poor quality and missing pages and issues were identified and recorded for the Sydney Morning Herald from 1831-1954. Approximately 7,500 pages need to be sourced from Australian institutions in hard copy and digitised in late 2009 to make the sequence complete.
- Discussions with Google took place so that Australian Newspapers can be harvested and made available via the Google News Archive service. See:
<http://news.google.com.au/archivesearch>
- Project documentation on the public website http://www.nla.gov.au/ndp/project_details was updated and new items added.
- Presentations on the ANDP were given at the Queensland Public Library Conference, the ANPlan/ANDP workshop, the 12th Australasian Congress of Genealogy and Heraldry and the NSLA workshop prior to the February NSLA meeting. One to one sessions for visitors from other national libraries (Indonesia, Singapore, Finland, Holland, New Zealand) and via teleconference were also given. Visits from representatives of State and Territory libraries as well as staff from Geoscience Australia and Australian War Memorial took place.

2. Work to be completed March 2009 – September 2009

- The ANDP team will commence working with the new panel of scanning and OCR contractors. It is anticipated that some of the Sydney Morning Herald will be OCR processed and made available via the Australian Newspapers beta service in this period.

- Development work will be undertaken on the Newspapers Content Management System to support more flexible ingest and management of content. In particular this work will enable working with multiple panel contractors and contributors.
- Development requirements and priorities for the Australian Newspapers beta service and the strategy for moving from beta to a production service will be determined.
- A communication channel will be set up for Australian Newspaper users to communicate with each other and the ANDP team.
- Contribution of content from the State Library of Victoria's newspapers digitisation activities will be tested and implemented.