

Australian Newspapers Digitisation Program



Search and Delivery Prototype – Design Brief

Author: Rose Holley (Manager - ANDP)

Version: 1.0

Date: 3 December 2007

Background

The Library has spent the last 7 months developing the storage and content management system for the ANDP service. This has been a challenging and resource intensive process. It was essential that the infrastructure was in place and operational before any development work could begin on the public search and delivery system, which is the last part of the software development process. The Library has decided to use the prototyping method of development for the search and delivery system. A prototype is an incomplete model of the future search and delivery system developed over a short period of time. The prototype will not have all the functionality or the 'look and feel' that the final system will have. The drivers for developing and releasing a prototype to stakeholders are that:

1. Stakeholders get the earliest opportunity to view an example system.
2. The Library gains feedback before signing off the final functional specification for the finished system.
3. The stakeholders are actively involved in the program.
4. The Library can continue to work in an 'agile' way that is cost effective.

This design brief outlines the key points about the development of the prototype. (Further information on prototyping is available here: http://en.wikipedia.org/wiki/Software_prototyping)

Data Content

'Pilot' data will be available in the prototype. Pilot data consists of 50,000 pages of content which were selected by State and Territory Libraries as representative sample of the 3 million page program. The pilot covers 12 newspaper titles spanning date ranges 1803 – 1945. It contains just over 8000 newspaper issues which equate to 50,000 individual pages and 500,000 individual articles. The pilot data is being processed using 2 external contractors (W & F Pascoe's for scanning work and Apex Co-Vantage for OCR, categorising and zoning). More information on the scope and coverage of the pilot data can be accessed here:

http://www.nla.gov.au/ndp/project_details/documents/ANDP_Pilot_Data.pdf

Goals

The overall goal of the program is to increase access to Australian newspapers. The search and delivery system will enable this by:

- Acting as a central point of access to all Australian newspapers
- Enabling full text searching across Australian newspapers
- Enabling browsing of Australian newspapers

The Library has reviewed other newspaper services and has a goal to offer a better service than any others by offering these things in particular:

- Faster searching and loading of pages
- Relevant results and easy to decide from the result list if an article is worth viewing
- Lots of ways of looking at things - newspaper, issue, page, article view
- Various options for viewing, zooming, moving around and printing
- Easy print option
- User interaction by ability to correct OCR, add annotations and tag articles.

Target Audience for finished service

The Library aims to make its collections accessible to *all* Australians. The newspaper service will not be targeted for a specific user group. However identified newspaper user groups are:

- Personal researchers and historians (including genealogists and local historians)
- Education sector (including school children, teachers, tertiary and higher education students)
- Public

Search and Delivery Requirements

It is essential that the final search and delivery system meets these requirements:

- Is simple, clear and easy to use
- Enables the user to quickly find what they are looking for
- Enables the user to view, navigate, print and save pages in multiple ways
- Enables the users to interact with the data and add value to the data
- Is accessible to visually impaired users
- Is supported by recent versions of the common internet browsers

Scope of Prototype

The newspapers prototype is targeted specifically at search and retrieval of only the pilot data from the Australian Newspapers Digitisation Program. It is out of scope to integrate search and retrieval of other resources simultaneously (e.g. Libraries Australia, Google, other online newspaper resources and indexes). Integration of these other resources will be addressed in version 1 of the Search and Delivery System.

Functionality

The purpose of the prototype is to prompt for feedback on functional requirements and features for the search and delivery system version 1. All functional requirements and features that have been identified so far are incorporated into a document here:

http://www.nla.gov.au/ndp/project_details/documents/ANDP_SearchDeliveryFunctionsFeatures.doc.

The document identifies those features which have been included in the prototype. Stakeholders can feedback on which features they think should or should not be included in the search and delivery system version 1, prioritise importance of features, and suggest new features for inclusion. The basic principle that has been applied in the prototype model is a move away from separate simple and advanced search screens and the user being forced to select search criteria as the first step. Instead the user is expected to perform a general search and then refine the results based on options given to them in the results screen. Most of this functionality is the same as that which would traditionally be given in a separate 'advanced search interface', however the mode of operation and search strategy are quite different.

Interface Design

During the 5 week prototype development a web designer will only be available for 9 days. Therefore limited interface design will take place. However the user should be able to navigate effectively and use the system. No colour/name or logo design will take place. Basic interface principles that will be applied are:

- Users will be able to search from anywhere on the site.
- There will be interfaces specific to particular ways of search and discover e.g. navigating through a newspaper, an issue, a page, an article (rather than having all discovery methods on a single interface).
- Keyword searching will be a 4 step process, with 4 interfaces –
 - Search
 - View results and select refinements/limits
 - View refined results
 - View specific items in results
- Interfaces will be clear to the users (without jargon, acronyms and excessive clutter/options)
- A clear distinction will be made between the results and options to refine the results.
- The interface will need to change when more data is added to the corpus. In the prototype it may be necessary to have visual representations of data contents and search examples since the pilot data is so limited.

Timescale

The Library will develop the prototype over a 5 week period. The prototype will be made available to the identified stakeholders in December 2007 and will be open for comment and feedback until mid January. (Exact dates will be supplied to stakeholders).

Stakeholders

For contractual reasons the prototype will only be accessible to stakeholders that have provided microfilm for digitisation. These stakeholders are State and Territory Libraries and W & F Pascoe's Ltd. The stakeholders are expected to provide feedback on the prototype. (Access details will be provided to stakeholders).

Branding/Name/Logo

The branding, naming and Logo of the newspaper service will not to be decided at this stage. The prototype will be released in the 'Library Labs' environment and will simply be called 'Newspapers Prototype'.

Feedback and next steps

Stakeholders can feedback by e-mail and telephone, informally or formally. The Library is not asking specific questions or conducting an online survey at this stage but is seeking more general feedback based around some general questions. It is expected that feedback on the prototype will be in these 3 broad areas:

- Functionality and features
- Look and feel (interface design)
- The pilot data.

(Feedback information will be supplied to stakeholders).

The Library will continue to develop the prototype and fix reported bugs during the release process. Feedback will be compiled and the results made transparent to stakeholders. The Library will take the feedback into consideration when developing the functional specification for the search and delivery system version 1. In an ongoing capacity the Library will continue to create and store content and further develop the infrastructure and content management system so that it supports a national collaboration model.