

Australian Newspapers Digitisation Program

Enhancements to consider for the Australian Newspapers Search and Delivery System

Author: Rose Holley (Manager - ANDP)

Version: 1.0

Date: 24 December 2008

A beta system was released to the public on 25 July 2008, specifically so that public users and Contributors could give feedback to assist the National Library with further development of the service and to ensure that it met user needs and expectations. Hundreds of suggestions, comments and feedback forms were received from the public between August and December 2008. Suggestions were also received from Contributors (State and Territory Libraries) and the ANDP team in this time. Many of the comments were overwhelmingly positive and these are recorded separately. The suggestions and comments have been collated and organised into groups by type. The items below are constructive suggestions for improvements to the search system that the ANDP team is in agreement with. Those marked with a priority of high and shaded in pink are considered by the ANDP team to be those items that should be addressed as soon as is practicable. Items have been given a high priority if many members of the public suggested the same thing (e.g. improvement to OCR correction); if the suggestion was made by the public, the ANDP team and also Contributors; or if the suggestion would considerably improve basic use of the newspapers (e.g. being able to login in easily, being able to search effectively).

The next steps are for the Library to determine when in 2009 further development will take place, how long it will take, and then when the beta system will be upgraded into Search and Delivery version 1 and officially launched. In the meantime the Australian Newspapers beta system will remain available to the public and new content will be added in the first quarter of 2009 as the Library works with new scanning and OCR contractors.

#	Enhancements grouped by type	Suggested by	Priority
1.	Searching newspapers		
1.1	Search for illustrations only	NDP	
1.2	Search for all papers published on a particular day and month without specifying a year (e.g. Birthday papers)	NDP	

1.3	Ability to search using geospatial techniques in a visual way (co-ordinates, by map/state).	NDP	
1.4	Ability to search using a timeline in a visual way.	NDP	
1.5	Ability to search across comments	NDP Public	High
1.6	Active help from authority files for Personal & Place names	NDP	
1.7	Ability to search across just tags (tag cloud is too big to browse).	NDP Public	High
1.8	Options to search across tags and text (and maybe comments too)	Public NDP	High
1.9	Option to search just corrected text only or in combination with all text		
1.10	Search on eras – e.g. Victorian, Great Depression	NDP	
1.11	Search within a title (by checkbox below search on title page).	NDP Public Contribs	High
1.12	Filter by state	Public	High
1.13	Filter by regional papers, state papers, national, broadsheets, tabloids, dailies, weeklies, Sunday papers at a later date when these are in	NDP	Low
1.14	Be able to specify exact search match only. At present system fuzzy searches and returns exact matches first followed by fuzzy results, using quotes does not change this.	Public	
1.15	More advanced search options e.g. proximity searching, fuzzy vs phrase, keeping restrictions on, within a title, comments, tags, OCR – anything else possible to get specific results and shorten results list. How to search for a / eg 2/32 battalion?	Public	High
1.16	Echo the search term in the search box (like Google, Worldcat) so users can see what they've searched on and change it	NDP K	High
1.17	How to exclude horses names from searching?	Public	
1.18	Filter on display and classified ads	Public	
1.19	On advanced screen put a 'search' button at top by search box (it is only at the bottom at present)	Public	High

1.20	To be able to search for personal names only in tags, and any other ways of searching more effectively for personal names?	Public	High
2.	Browsing newspapers		
2.1	Browse back and forward through multiple newspaper titles on a single date continuously	NDP	
3.	Results List		
3.1	Show which articles have been corrected in results list with a icon or something beside	Public	High
3.2	Show which articles have been tagged in results list	Public	High
3.3	Display some of the tags of an article in results list if it has been tagged	NDP	High
3.4	Related items to your search ...	NDP	
3.5	Related articles to your search....		
3.6	Mark items and re-display	NDP	
3.7	Change number of results displayed	NDP Public	High
3.8	Display thumbnail images of pages	NDP	Low Remove?
3.9	Display image of headline	NDP	Low Remove?
3.10	Next' buttons at bottom of results list are not going to next page of result but to next 10 pages. Need next page, first page, last page results jump or to number. Buttons not working as users expect.	Public	High
3.11	Sort results bug	NDP	
4.	Article Level View		
4.1	E-mail article (or just raw text??)	NDP	
4.2	Save article button / save to 'bookbag'	NDP	
4.3	View related articles	NDP	Medium
4.4	Cite this article in different citation formats	NDP	Low

4.5	When you do a keyword search and arrive at article level can you use the Tab key to jump to the next highlighted word in article rather than having to scroll down?	Public	
4.6	Navigation of long multiple page articles difficult since links are only at top of column/text	NDP Public	
4.7	Users bookmarking articles in progress which then disappear. (page has a pi but articles do not until accepted hence appear to disappear when accepted). How will users know and find them again?	Public	
5.	Page Level View		
5.1	Cite this page in different citation formats	NDP	Low
5.2	E-mail this page	NDP	
5.3	Rotate page (for tables only, all pages will be appearing correct way up already).	NDP	
6.	Issue Level View <i>This was available in prototype but is not available in Beta. It is desirable to have issue level view when further development takes place.</i>		
6.1	View a list of articles by page.	NDP	
6.2	Persistently cite this issue	NDP	
6.3	Follow links to a specific article.	NDP	
6.4	Follow links to a specific page	NDP	
6.5	Cite this issue in different citation formats	NDP	
6.6	Navigate back/forward to previous/next issue	NDP	
6.7	Jump to specified page in issue	NDP	
6.8	E-mail this issue	NDP	
6.9	View thumbnails for entire issue (without contents)	NDP	
6.10	View, save and Print a PDF version of the issue (sized to print on an A3 page without adjustment), with citation	NDP	
6.11	Search within this Issue	NDP	

7.	User Interaction (at article level)		
	Text Correction		
7.1	To be able to do OCR correction more easily and quicker, power user mode, especially for long blocks of text	NDP Public	High
7.2	Be able to add missing lines in OCR text (user found a workaround to do this adding a space at a time and saving) quite a few articles have last 2 lines missing. Adverts have no lines for text.	NDP Public	High
7.3	Want to know overall stats for text correction. Top 10 and their own is not enough, want the big picture and where they are in complete list, and how we are doing.	Public	
7.4	Text correctors to be fed articles to correct rather than doing it randomly (for correctors who want to just correct and do a certain amount per week), or to be able to see/search for articles not yet corrected.	Public	Medium?
7.5	Help fix this text keeps hovering and gets added into the text correction sometimes	NDP Public	High
7.6	Some users to be volunteer moderators for OCR corrections, tags, comments	Public	High
7.7	Want acknowledgement/thanks/publicity/certificates for text correctors to keep motivation up and encourage other people to join in.	Public	
7.8	Correct re-keyed fields if incorrect (title, subtitles, author, abstract)	NDP	
7.9	Roll back and reporting OCR changes (e.g. we and user see number of their corrections)	NDP	
7.10	Be able to make global edits, report global edits needed or have post processing done on OCR text before loaded e.g. winks and spirits to wines and spirits, streefc to street.	Public	
7.11	If text is hard to read/correct want to flag this so other users can give second opinion.	Public	
7.12	The message about time delay of updating text should only appear once in a session, not every time you save.	Public	
7.12a	How to input symbols into text correction eg pound and em dash? Also you don't seem to be able to specify italics eg for ship names.	Public	
7.12b	Difficult to navigate to correct text in multipage articles.	Public	
7.12c	Save changes button is at top of screen and on long article have to go all way to top again to save, can it be at bottom as well or near text correction. It could also be a green tick instead of save.	Public	

7.12d	How to more easily copy into a doc the ocr text only.	Public	
	Communication and sharing		
7.13	Further develop user profiles. Allow user to add photo, give research interests, have open or closed profile, let other users contact them via profile (similar to profiles on family tree forums). Give user better stats in their profile and if correctors where they are in complete ranking. Many users haven't noticed they have a profile since it is hidden (click on your name on home page).	NDP Public	High
7.14	To be able to have a profile/login for a group (top tag is being used by group LRSA), or can indiv users join a group to identify research interests/joint research.	NDP	medium
7.15	To build and connect into the text correctors community (it appears there are about 1500 people in this so far and some doing only this, not tagging as well).		
7.16	Be able to communicate with other users	Public	High
7.17	Linked Forum set up by ANDP	NDP Public	
7.18	RSS/alerting feeds for new content (pages, titles, date ranges, numbers of each)	NDP Public Contribs	High
7.19	Users creating and publicly sharing trails/topic features/timelines	NDP	
	Tags		
7.20	Add tags (keywords) to Pages and Issues, and follow links to other pages/issues/articles with the same tags	NDP	
7.21	Edit other peoples tags (or moderate?)	NDP Public	Medium
7.22	Can tags be broken into personal names, place names, events or given a property of this and will this then link to searching tags of these types? E.g. to be able to search on 'scotland' as a person not a place.	Public	?
7.23	Tag where events happened on map (linked to geospatial visual searching)	NDP	
7.24	Add tags/comments to a specific place in the article (for long articles). E.g want to pinpoint name in classified ad, family announcement	NDP Public	
7.25	Make suggestions when people try to add tags for synonyms e.g. WWI use first world war etc	NDP	
7.26	Can names be standardised in the existing 18,000 tags?	Public	

7.27	To be able to print own tags and saved articles	Public	
7.28	Tags are only 30 characters. Too short to enter the name of a parliament bill. Can it be lengthened?	Public	
7.29	To be able to Geotag using co-ordinates in tag.	Public	
	Comments		
7.30	Edit or delete your comments	NDP Public	High
7.31	Manage your comments	NDP	
7.32	Browse all comments	NDP Public	High
7.33	Want to paste url's into comments field	Public	
	Tracking Research		
7.34	How can users keep track of their research? (users appear to be using tags for this purpose and there could be a lot better ways?)	NDP	High
7.35	To be able to bookmark pages, to classify them, to associate notes with them and also to assert typed relationships (predicates) between bookmarked pages. Going one step further, users could create their own subjects/objects e.g. "Wilhelm Wenceslas Engel" who would have certain properties like birth and death dates, places of birth and death (and so on).	People Australia	
7.36	When I am doing research and searching on the same/different term how can I know what articles I have previously looked at? I am wasting a lot of time re-reading the same articles. Could the hyper link show in a different colour once I've viewed it like Google and can this be retained for my next session?	Public	
7.37	Build relations between articles e.g. the saga of the Polar bear at Melbourne zoo	NDP	
7.38	Saving searches – see also authentication	NDP	
7.39	Search History – see also authentication	NDP	
7.40	Genealogists able to add in family trees/research lines somehow	NDP	
7.41	Want to be able to keep a track of everything they have done indefinitely eg user profile to not just show last 10 things but option to view by month/week/year number of corrections/tags or ALL.	Public	

	Other		
7.42	Personalisation of home page – see also authentication		
7.43	User creates substantial/significant/long new information e.g. “What happened next” to follow on from news articles	NDP	
7.44	Allow ‘mashing’ of data (re-purposing)	NDP	
7.45	Edit category of article if it is wrong	NDP	
7.46	Competition to create publicity youtube video for ANDP/video clips to help users do things eg OCR correction (see help below)	Public	
8.	Newspaper Title Page View		
8.1	Follow links to previous, subsequent and related titles for this newspaper title (e.g. Brisbane Courier to Courier Mail)	NDP Public Contribs	High
8.2	Know what newspapers are in a ‘series group’	NDP Public	High
8.3	Newspaper titles to have consistency with Voyager, Libs Australia (and titles exported to Google see harvesting section) Voyager records may need to be altered and made correct..	NDP	High
9.	User Interface		
9.1	Highlight recent text corrections – home page	NDP	
9.2	Display note that we are aware of bad pages next to page image e.g. fragmented ones	NDP	
9.3	Feedback contact NDP/ANPLAN for found missing pages, and generate lists on demand for known missing pages.	NDP	
9.4	Display date microfilm was created	NDP	
9.5	Font type and size fixed	NDP	
9.6	Acknowledge sponsorship of digitisation	NDP	
9.7	Users still finding it hard to understand/locate saving and printing. Especially difficulties with printing as small part of a long page classified ad. There are save and print buttons (which sometimes do same thing depending what level you are at) and also the right mouse click to save to jpg which only the computer literate find by themselves. Still a lot of people going File-Save which does not work...	NDP Public	High

9.8	Usability on small screens?	NDP	
9.9	Remove 'beta' from interface when it goes to v1 and is launched	NDP	
9.10	Remove feedback tab and survey at some point	NDP	
9.11	Have 'contact us' tab when go to v1 and there is someone to support service	NDP	
9.12	Radio buttons 'in these results' and 'new search' clutter search area. Button 'in this title' was never added in but is needed. How does 'refine search' on left hand side relate to 'in these results' on right hand side?	NDP K	
9.13	How would users find out/know that double clicking an image will zoom and centre it?	NDP	
9.14	Not sure zoom levels are being retained as they should?	NDP Public	
10.	Reliability, Speed and Accessibility		
10.1	System designed to be usable on mobile devices	NDP	
10.2	System accessible to vision impaired (or files available to download in daisy or whole database, or text only version?)	NDP	
10.3	Interface functions across latest versions of multiple browsers	NDP	
10.4	Want time delay on updating OCR corrections to be reduced so they don't accidentally correct something twice.	Public	
10.5	System is not accessible for 1.5 hrs at Australian eastern time 12pm (WA =10pm, UK=1pm).	NDP Public	High
11.	Authentication and Personalisation		
11.1	Authentication that enables personalisation of home page, saving searches and articles	NDP Public	
11.2	Authentication for moderation and management (e.g user moderators and Library)	NDP	
12.	Integration of Other Content (out of scope for Beta, but desirable for v1)		
12.1	Searching across other free digital collections and resources e.g. Libraries Australia, Newspapers, Google	NDP	
12.2	Integrating hard copy and online newspaper indexes e.g. Argus Online	NDP	Low
12.3	Integrating other full text resources e.g. books and journals	NDP	

12.4	Integrating/searching across subscription only newspaper resources e.g. SMH, ProQuest Historical Newspapers	NDP	
12.5	Use of LCSH, DDC or topics/subjects tags in broader sense.	NDP	
13.	High quality Image Ordering		
13.1	Order high quality tiff image of pages and issues	NDP Contribs	Medium
14.	Legal		
14.1	Modify and merge disclaimer/terms of use. Add item that pages may be removed and replaced and therefore user data may be lost.	Legal adv	?
14.2	Put a cultural sensitivity statement in the disclaimer re picture and also tone of articles re aboriginals	Public	
15.	Help, about and guidance		
15.1	Context sensitive help	NDP Public	Medium
15.2	Video clips to demo text correction techniques, tags, comments	NDP Public	Medium
15.3	Broad based FAQ (FAQ getting too long on about page)	Public	
15.4	Guidelines for text correctors e.g. italics, em dashes, pound symbol. Lots unsure how to treat hyphens in correcting and how to put in symbols.	Public	
15.5	Guidelines for taggers e.g how to enter personal names	Public	
15.6	Improve the graph showing content on the about page. Users still having trouble getting an idea of scope of data overall.	Public	High
15.7	Want a page within the system to show new titles, number of pages, date ranges added to system, per week and also number of articles corrected per week (see also RSS alerts for this in user interaction section).	Public	High
15.8	More examples of searching in help and syntax if there is any and how to deal with punctuation	Public contribs	High
16.	Login (Some of these are for User Management System Development)		
16.1	If you find an article then want to correct it, then login you are taken to home page not direct back to article. Want to come to article direct after logging in.	Public	High

16.2	Not clear when or if you need to login – how to address this?	Public	
16.3	Difficulty with capture – too difficult may have to try several times (lots of users)	Public	
16.4	Several users say they can't see the capture and then the system freezes	Public	
16.5	Using symbols in password doesn't work and it should (there are no rules for passwords showing)	Public	
16.6	Difficulties with successful sign up process (lots of users). Not clear if you haven't completed registration process. Also if you use symbols in your password you get an error message saying the user ID is in use but this is not right. Users change their userid several times but keep the same password with symbols and so can't sign up...	Public	High
16.7	User id is already taken – then what? (This message is being used if password has symbols in it so is very misleading and in that case changing the userid still doesn't work).	Public	High
16.8	Cancel stale userid creation and password change after 3 days	NDP	
16.9	Alert in e-mail confirmation if another user has same e-mail	NDP	
16.10	What is timeout and can it be lengthened. If read articles too long you get timed out.	Public	
17.	Harvesting/download of all data		
17.1	Harvest site for Google – need to have agreed newspaper title in here, and should be same as search system, libs Australia – needs more cataloguing discussion	NDP	
17.2	Should we have site where orgs can download all the data?	NDP	

Browsers on which beta works:

	Win 98	Win 2000	Win XP	Win Vista	Mac 10.3.x	Mac 10.4
IE 7.0			A-grade	A-grade		
IE 6.0	A-grade	A-grade	A-grade			
Firefox 2.	A-grade	A-grade	A-grade	A-grade	A-grade	A-grade
Firefox 1.5.	A-grade	A-grade	A-grade		A-grade	A-grade
Opera 9.	A-grade	A-grade	A-grade		A-grade	A-grade
Safari 2.0						A-grade