

# Australian Newspapers Search & Delivery System Functional Specification

**Australian Newspapers:** Version 1  
**Production release:** August 2009  
**Spec written by:** Bronwyn Lee  
**Date:** 14 August 2009

## Introduction

A prototype search & delivery system containing data from the 50,000 page pilot was released on 10 December 2007 and access was limited to State and Territory Library stakeholders for contractual reasons because the data at that time did not reach NLA's quality standards.

Australian Newspapers Beta was made available to the public unannounced on 25 July 2008. It was named Beta since on release the data was limited to the pilot data, which had been re-processed to meet the quality standards, plus any extra data processed by Apex and accepted by NLA up till then. Extra data continued to be added as it was processed and accepted.

A number of improvements were made until, on 6 August 2009, a full production service was released as Australian Newspapers v1.0. Post production fixes were applied in the following week. This document describes the system as it was on 14 August 2009.

## List of use cases and interface objects

### By interface object

<i>Interface object</i>	<i>Use case</i>
<a href="#">Standard Header</a>	<a href="#">Search for articles</a>
<a href="#">Home</a>	<a href="#">Login</a> <a href="#">Search for articles</a> <a href="#">Browse newspaper issues</a>
<a href="#">About</a>	<a href="#">Other: Website information ('static' pages)</a>
<a href="#">Help</a>	<a href="#">Other: Website information ('static' pages)</a>
<a href="#">Login</a>	<a href="#">Login</a>
<a href="#">Signup</a>	<a href="#">Signup</a>
<a href="#">Terms of Use</a>	<a href="#">Other: Website information ('static' pages)</a>
<a href="#">Reset Forgotten Password</a>	<a href="#">Reset password</a>
<a href="#">Browse</a>	<a href="#">Browse newspaper issues</a>
<a href="#">Newspaper Titles List</a>	<a href="#">Browse newspaper issues</a>
<a href="#">Issues by Date List</a>	<a href="#">Browse newspaper issues</a>
<a href="#">Tagged Articles List</a>	<a href="#">View tags</a>
<a href="#">Tag Cloud</a>	<a href="#">View tags</a>
<a href="#">Search Results</a>	<a href="#">Search for articles</a>
<a href="#">- No Match</a>	<a href="#">Search for articles</a>
<a href="#">Advanced Search</a>	<a href="#">Search for articles</a>
<a href="#">Newspaper View</a>	<a href="#">View a newspaper</a> <a href="#">Browse newspaper issues</a>
<a href="#">Page View</a>	<a href="#">View a page</a>

	<a href="#">View an issue</a>
<a href="#">Save Page as PDF</a>	<a href="#">Save / print a page</a>
<a href="#">Article View</a>	<a href="#">View an article</a>
- <a href="#">Edit Immediately or Become a Member</a>	<a href="#">Add / edit a comment</a> <a href="#">Add / edit tags</a> <a href="#">Correct the text</a>
- <a href="#">Add/Edit Tags</a>	<a href="#">Add / edit tags</a>
- <a href="#">Add / Edit Comment</a>	<a href="#">Add / edit a comment</a>
- <a href="#">Show Comments</a>	<a href="#">View comments</a>
- <a href="#">Correct the Text</a>	<a href="#">Correct the text</a>
<a href="#">Print Article</a>	<a href="#">Save / print an article</a>
<a href="#">Save Article as PDF</a>	<a href="#">Save / print an article</a>
<a href="#">Save Article as Image</a>	<a href="#">Save / print an article</a>
<a href="#">Comment History</a>	<a href="#">View comments</a>
<a href="#">Profile for User</a>	<a href="#">View tags</a>
<a href="#">Show Corrections</a>	<a href="#">View text corrections</a>
<a href="#">Your Comments</a>	<a href="#">View comments</a>

#### By use case

<i>Use case</i>	<i>Interface object</i>
<a href="#">Signup</a>	<a href="#">Signup</a>
<a href="#">Login</a>	<a href="#">Login</a> <a href="#">Home</a>
<a href="#">Logout</a>	-
<a href="#">Reset password</a>	<a href="#">Reset Forgotten Password</a>
<a href="#">Browse newspaper issues</a>	<a href="#">Home</a> <a href="#">Browse</a> <a href="#">Newspaper Titles List</a> <a href="#">Issues by Date List</a> <a href="#">Newspaper View</a>
<a href="#">View a newspaper</a>	<a href="#">Newspaper View</a>
<a href="#">View an issue</a>	<a href="#">Page View</a>
<a href="#">View a page</a>	<a href="#">Page View</a>
<a href="#">View an article</a>	<a href="#">Article View</a>
<a href="#">Search for articles</a>	<a href="#">Standard Header</a> <a href="#">Home</a> <a href="#">Search Results</a> <a href="#">No Match</a>
<a href="#">View comments</a>	<a href="#">Show Comments</a> <a href="#">Your Comments</a> <a href="#">Comment History</a>
<a href="#">View text corrections</a>	<a href="#">Show Corrections</a>
<a href="#">Save / print an article</a>	<a href="#">Print Article</a> <a href="#">Save Article as PDF</a> <a href="#">Save Article as Image</a>
<a href="#">Save / print a page</a>	<a href="#">Save Page as PDF</a>
<a href="#">Add / edit a comment</a>	<a href="#">Edit Immediately or Become a Member</a> <a href="#">Add / Edit Comment</a>
<a href="#">Add / edit tags</a>	<a href="#">Edit Immediately or Become a Member</a> <a href="#">Add/Edit Tags</a>

<a href="#">Correct the text</a>	<a href="#">Edit Immediately or Become a Member</a> <a href="#">Correct the Text</a>
<a href="#">View tags</a>	<a href="#">Tag Cloud</a> <a href="#">Tagged Articles List</a>
<a href="#">View tags</a>	<a href="#">Profile for User</a>
<a href="#">Other: Website information ('static' pages)</a>	<a href="#">Terms of Use</a> <a href="#">About</a> <a href="#">Help</a>
<a href="#">Other: Statistics</a>	-

## Use Cases

### Use case: Sign up

Actors	User
Description	This use case allows a user to sign up as a registered user.
Preconditions	None. The user may sign up when entering the system (the Home page) or when about to perform an action which may require or be advantageous to the user that the user be identified.
Postconditions	The user has a registered user name, password and email address recorded in the system.
Basic Course	<ol style="list-style-type: none"> <li>1. System presents a signup form requesting the user's details.</li> <li>2. System also requires the user complete a task to prove the form has not been populated by an automated process (captcha challenge)</li> <li>3. System also requires user to accept terms and conditions of use.</li> <li>4. User enters their details, completes the required tasks and submits the form.</li> <li>5. System checks user input is valid.</li> <li>6. System records user in user management system and sends message to the user's email address with a url link to complete registration.</li> <li>7. System presents message to user asking them to complete their registration by clicking on the url link in the email.</li> <li>8. User opens the email message and clicks on the link.</li> <li>9. System completes the user registration process in the user management system.</li> </ol>
Alternative Courses	
Exceptions	<p>At step 5, system finds some of the user input is not valid.</p> <ul style="list-style-type: none"> <li>• System presents error message</li> </ul> <p>At step 8 user does not respond to email.</p> <ul style="list-style-type: none"> <li>• After a period of time (3 days) the user registration becomes 'stale' and user name and password creation is cancelled</li> <li>• If user invokes the url in the email after 3 days, system presents an explanatory message.</li> </ul>
Implementation Notes	The NDP system uses a generic User Management System which is shared currently with the Digital Preservation workflow system. The UMS is documented at <a href="https://wiki.nla.gov.au/display/SOA/User+Management+Service">https://wiki.nla.gov.au/display/SOA/User+Management+Service</a> (see the attachment to this wiki page for documentation which is relatively non-technical).

Interface Object/s	<a href="#">Signup</a>

Back to [List of use cases and interface objects](#)

### Use case: Login

Actors	User
Description	This use case allows users to log into the system. Login is optional.
Preconditions	The user is not already logged in. The user has either just entered the system or the user is already using the system and wants to perform an action where login is desirable or required.
Postconditions	The user is logged into the system and all functionality the user is authorised for is available.
Basic Course	<ol style="list-style-type: none"> <li>1. System presents a login form requesting the user's username and password.</li> <li>2. User enters username and password and elects to continue.</li> <li>3. System authenticates the user .</li> </ol>
Alternative Courses	At step 2 the user declines to log in.
Exceptions	<p>Continues after step 2: System does not recognise the username or the username and password combination is incorrect.</p> <ul style="list-style-type: none"> <li>• System displays appropriate message.</li> <li>• System starts basic course again.</li> </ul>
Implementation Notes	<ul style="list-style-type: none"> <li>• Login is optional. All functionality will be available to anonymous users.</li> <li>• If users wish to have comments, tags and text corrections credited to them, or if they wish to add private comments or tags (i.e. not visible to other users), they need to log in.</li> <li>• If not already logged in, users will be prompted to log in when they go to add comments, tags or OCR corrections.</li> <li>• The NDP system will pass user details to the NLA User Management System for authentication.</li> </ul>
Interface Object/s	<a href="#">Login</a> <a href="#">Home</a>

Back to [List of use cases and interface objects](#)

### Use case: Reset password

Actors	User
Description	Allows users who have forgotten their password to reset it.
Preconditions	User is trying to log in.
Postconditions	User is able to log in.
Basic Course	<ol style="list-style-type: none"> <li>1. System presents a Reset Password form.</li> <li>2. User enters user name, proves they are human (by completing a captcha challenge) and submits the form.</li> <li>3. System validates user input, looks up user in the User Management System (UMS) and finds their email address.</li> </ol>

	<ol style="list-style-type: none"> <li>4. System sends email message to the user's email address.</li> <li>5. System presents message asking user to check their email and to follow the instructions in the email message.</li> <li>6. User opens email and follows instructions i.e. clicks on link in email or pastes the url into their browser.</li> <li>7. System presents a Reset password form which allows user to reset their password.</li> <li>8. User enters a password, then enters it again, and submits the form.</li> <li>9. System validates the user's input (e.g. password not repeated correctly).</li> <li>10. System presents a confirmation message that the password has been successfully changed and invites the user to log in.</li> </ol>
Alternative Courses	
Exceptions	<p>User input is invalid (.e.g user name doesn't exist, captcha incorrect, password not repeated correctly)</p> <ul style="list-style-type: none"> <li>• System presents error message with the form.</li> <li>• User re-enters details in form.</li> </ul> <p>At step 6 user does not respond to email within 3 days:</p> <ul style="list-style-type: none"> <li>• After a period of time (3 days) the user registration becomes 'stale' and password reset is cancelled</li> <li>• If user invokes the url after 3 days, system presents an explanatory message.</li> </ul>
Implementation Notes	The password is stored hashed, i.e. NLA staff cannot see what the password is when they view user details in the UMS.
Interface Object/s	<a href="#">Reset Forgotten Password</a>

Back to [List of use cases and interface objects](#)

### Use case: Log out

Actors	User
Description	
Preconditions	The user is already logged in.
Postconditions	The user's session is terminated. The user does not have access to functions requiring login.
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to log out.</li> <li>2. System terminates the user's session.</li> <li>3. User is returned to the Home page.</li> </ol>
Alternative Courses	
Exceptions	•
Implementation Notes	<p>Kent 25/07/08: Ive changed the logoff code to delete all cookies it finds.</p> <p>Ive also changed "home" to check that if there is no session user, then if there is a newspaperUser cookie, to make sure it is cleared.</p> <p>The "newspaperCaptcha" cookie is trickier to deal with - there's no "logout" to nullify this; I have a vague feeling it might be better to not have these cookies in article, but instead to represent them somehow</p>

	in the jsp session, and set some javascript variables to be tested by any client side code (to see whether the client thinks it is in a session with a valid user id or a capture) - I guess I'm slightly worried that a rogue client could just present some cookies with the username they wanted to - maybe the updating code checks for this (I havent checked)
Interface Object/s	

Back to [List of use cases and interface objects](#)

### Use case: Browse newspaper issues

Actors	User
Description	This use case allows users to select a date and see what issues are available for any newspaper, or to select a newspaper and see what issues are available for that newspaper.
Preconditions	-
Postconditions	The user is able to view an issue of a newspaper.
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to browse by date.</li> <li>2. System presents a browse by date form.</li> <li>3. User selects a year.</li> <li>4. User selects a month.</li> <li>5. System displays dates for which issues are available.</li> <li>6. User selects an issue to view.</li> <li>7. System displays the first page of the selected issue.</li> </ol>
Alternative Courses	<p>User chooses to browse newspaper titles.</p> <ul style="list-style-type: none"> <li>• User selects a newspaper title.</li> <li>• System displays the selected newspaper with a browse by date form containing only dates for that newspaper.</li> </ul> <p>Resume Basic Course at step 3.</p>
Exceptions	
Implementation Notes	Users can select from a list of all newspaper titles or can choose a state first, to select from a list of newspapers from that state.
Interface Object/s	<a href="#">Home</a> <a href="#">Browse</a> <a href="#">Newspaper Titles List</a> <a href="#">Issues by Date List</a> <a href="#">Newspaper View</a>

Back to [List of use cases and interface objects](#)

### Use case: View a newspaper

Actors	User
Description	This use case allows users to view details of a newspaper.
Preconditions	-
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User selects a newspaper to view.</li> <li>2. System displays the selected newspaper.</li> </ol>

Alternative Courses	
Exceptions	
Implementation Notes	The newspaper title is hyperlinked to the Newspaper View in the 'coverage' graph on the About screen and the 'about' link after the newspaper title in Article View and Page View. It is not hyperlinked in the Search Results as the focus is on navigating to articles there.
Interface Object/s	<a href="#">Newspaper View</a>

Back to [List of use cases and interface objects](#)

### Use case: View an issue

Actors	User
Description	This use case allows users to view details of an issue (an issue is defined by a newspaper and a date).
Preconditions	-
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User selects an issue to view.</li> <li>2. System displays the first page of the selected issue.</li> </ol>
Alternative Courses	
Exceptions	
Implementation Notes	Users can link to the first page of the issue in some places where it is cited e.g. in the 'on this day' article on the home page; also in the issue, page and article views through next/previous issue links. It is not hyperlinked in the Search Results as the focus is on navigating to articles there.
Interface Object/s	The <a href="#">Issue View</a> is now the same as the Page View of the first page of the selected issue. The Issue View where users could see thumbnails of all pages in the issue at once was in the ndp prototype but this was removed in the Oxide redesign which integrated the article, page and issue views.

Back to [List of use cases and interface objects](#)

### Use case: View a page

Actors	User
Description	This use case allows users to view details of a page in an issue of a newspaper.
Preconditions	-
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User selects a page to view.</li> <li>2. System displays the selected page.</li> </ol>
Alternative Courses	
Exceptions	
Implementation Notes	Users can link to Page View in the issue, page and article views through next/previous page links and also in the article view through

	the 'view entire page' link. It is not hyperlinked in the Search Results as the focus is on navigating to articles there.
Interface Object/s	<a href="#">Page View</a>

Back to [List of use cases and interface objects](#)

### Use case: View an article

Actors	User
Description	This use case allows users to view details of an article in an issue of a newspaper. An article may occur on multiple pages in an issue but an article 'continuation' in another issue is regarded as a separate article.
Preconditions	-
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User selects an article to view.</li> <li>2. System displays the selected article.</li> </ol>
Alternative Courses	
Extensions	<p>Use cases:</p> <ul style="list-style-type: none"> <li>• View comments</li> <li>• View text corrections</li> <li>• Save/print an article</li> <li>• Add/edit a comment</li> <li>• Add/edit tags</li> <li>• Correct the text</li> </ul>
Exceptions	
Implementation Notes	Users can link to Article View from the 'on this day' article on the home page; in the issue view; the page view; tagged article list, and search results.
Interface Object/s	<a href="#">Article View</a>

Back to [List of use cases and interface objects](#)

### Use case: Search for articles

Actors	User
Description	This use case allows users to search for articles containing certain words or which meet certain other criteria.
Preconditions	-
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User enters search criteria and submits the search.</li> <li>2. System parses the user input and executes the search.</li> <li>3. System presents the search results i.e. a list of articles meeting the search criteria.</li> <li>4. User selects an article to view in more detail.</li> <li>5. System displays the article.</li> </ol>
Alternative Courses	
Exceptions	<p>User does not enter any search terms.</p> <ul style="list-style-type: none"> <li>• System retrieves all articles.</li> </ul>

	<p>No articles meet the search criteria.</p> <ul style="list-style-type: none"> <li>System presents a 'no match' message with some suggestions for trying again.</li> </ul>
Implementation Notes	<p>Search rules and functionality is documented in Help. The metadata supplied by NLA and Apex, the full text supplied by Apex, and corrections to the text are searchable. Tags and annotations are not searchable at present.</p>
Interface Object/s	<p><a href="#">Standard Header</a>  <a href="#">Home</a>  <a href="#">Search Results</a>  <a href="#">No Match</a></p>

Back to [List of use cases and interface objects](#)

### Use case: View text corrections

Actors	User
Description	This use case allows users to view the corrections that have been made to the OCR text
Preconditions	
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>User chooses to view OCR corrections which have been made to an article.</li> <li>System displays all the OCR corrections for the article.</li> </ol>
Alternative Courses	
Exceptions	
Implementation Notes	
Interface Object/s	<a href="#">Show Corrections</a>

Back to [List of use cases and interface objects](#)

### Use case: Save/print an article

Actors	User
Description	This use case allows users to save (download) or print an article.
Preconditions	This use case extends 'View an article' use case.
Postconditions	User has a paper copy of the article or user has a file copy of the article which can be printed or used later.
Basic Course	<ol style="list-style-type: none"> <li>User chooses to save or print the article.</li> <li>User selects a file format if more than one format is available.</li> <li>System presents the article in the selected file format.</li> </ol>
Alternative Courses	
Exceptions	
Implementation Notes	Both JPEG and PDF formats are offered.

Interface Object/s	<a href="#">Print Article</a> <a href="#">Save Article as PDF</a> <a href="#">Save Article as Image</a>

Back to [List of use cases and interface objects](#)

### Use case: Save/print a page

Actors	User
Description	This use case allows users to save (download) with or without printing a page.
Preconditions	This use case extends 'View a page' use case.
Postconditions	User has a paper copy of the page or user has a file copy of the page which can be printed or used later.
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to save the page.</li> <li>2. User selects a file format if more than one format is available.</li> <li>3. System presents the page in the selected file format.</li> <li>4. User may use the browser or reader software on their computer to print the page.</li> </ol>
Alternative Courses	
Exceptions	
Implementation Notes	PDF is the only format offered.
Interface Object/s	<a href="#">Save Page as PDF</a>

Back to [List of use cases and interface objects](#)

### Use case: Add/edit a comment

Actors	User
Description	This use case allows users to add a new comment or edit (revise) an existing comment.
Preconditions	This use case extends 'View an article' use case.
Postconditions	Users can view the new or changed comment.
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to add a new comment.</li> <li>2. System presents a form asking the user to prove they are human (captcha challenge), or the user is prompted to log in.</li> <li>3. User proves they are human.</li> <li>4. System presents a form where the user may enter their comment.</li> <li>5. User inputs their comment and may choose not to make the comment public.</li> <li>6. User submits the form.</li> <li>7. System checks the user's input against validation rules.</li> <li>8. System saves the comment.</li> <li>9. System presents message confirming that the comment has been saved.</li> </ol>
Alternative Courses	User chooses to edit a comment. Users can edit any comment they are able to view (i.e. comments they themselves have created or

	<p>public comments created by anyone else).</p> <ul style="list-style-type: none"> <li>• At step 4 system presents form populated with latest version of the comment.</li> <li>• At step 5 user edits the comment.</li> <li>• Resume basic course at step 6.</li> <li>• At step 8 system saves the new version of the comment, while keeping all the older versions of the comment.</li> </ul>
Exceptions	<p>Exception 1) If the user is already logged in, omit step 2.</p> <p>Exception 2) At step 7: System finds part/s of the user input are not valid. Alt step 8: System silently removes or changes the invalid part/s (to make them valid) and saves only the valid part/s of the user input.</p> <p>Exception 3) Alt step 3 or 6: User chooses to cancel. Alt step 4 or 7: Use case ends. User input is not saved.</p>
Implementation Notes	.
Interface Object/s	<a href="#">Edit Immediately or Become a Member</a> <a href="#">Add / Edit Comment</a>

Back to [List of use cases and interface objects](#)

### Use case: Add/edit tags

Actors	User
Description	This use case allows users to add tags to an article. Users may also change or delete tags they have created.
Preconditions	This use case extends 'View an article' use case.
Postconditions	Users can view the new or changed tag/s.
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to add or edit tags.</li> <li>2. System presents a form asking the user to prove they are human (captcha challenge), or the user is prompted to log in.</li> <li>3. User proves they are human.</li> <li>4. System presents a form populated with any tag/s the user has already assigned to the article.</li> <li>5. User adds, edits or deletes tags and may choose not to make the tags public.</li> <li>6. User submits the form.</li> <li>7. System checks the user's input against validation rules.</li> <li>8. System saves the tags.</li> <li>9. System presents message confirming that the tags have been saved.</li> </ol>
Alternative Courses	
Exceptions	<p>Exception 1) If the user is already logged in, omit step 2.</p> <p>Exception 2) At step 7: System finds part/s of the user input are not valid. Alt step 8: System silently removes or changes the invalid part/s (to make them valid) and saves only the valid part/s of the user input.</p> <p>Exception 3) Alt step 3 or 6: User chooses to cancel.</p>

	Alt step 4 or 7: Use case ends. User input is not saved.
Implementation Notes	
Interface Object/s	<a href="#">Edit Immediately or Become a Member</a> <a href="#">Edit Tags</a>

Back to [List of use cases and interface objects](#)

### Use case: Correct OCR text

Actors	User
Description	This use case allows users to correct i.e. make changes to the text which may be the original OCR'd text or text which has been previously corrected. Includes adding missing text.
Preconditions	This use case extends 'View an article' use case.
Postconditions	Users can view the changed text and it also becomes searchable.
Basic Course	<ol style="list-style-type: none"> <li>1. User selects a line of text to change.</li> <li>2. System presents a form asking the user to prove they are human (captcha challenge), or the user is prompted to log in.</li> <li>3. User proves they are human.</li> <li>4. System presents a form where the user may enter their changes.</li> <li>5. User inputs their changes.</li> <li>6. User submits the changes.</li> <li>7. System checks the user's input against validation rules.</li> <li>8. System saves the changes and indexes the changed text for searching (any older versions and the original text are kept).</li> <li>9. System presents message confirming that the changes have been saved.</li> </ol>
Alternative Courses	
Exceptions	<p>Exception 1) If the user is already logged in, omit step 2.</p> <p>Exception 2) At step 8: System finds part/s of the user input are not valid. Alt step 9: System silently removes or changes the invalid part/s (to make them valid) and saves only the valid part/s of the user input.</p> <p>Exception 3) Alt step 3 : User chooses to cancel. Alt step 4: Use case ends.</p>
Implementation Notes	
Interface Object/s	<a href="#">Edit Immediately or Become a Member</a> <a href="#">Correct the Text</a>

Back to [List of use cases and interface objects](#)

### Use case: View tags

Actors	User
--------	------

Description	This use case allows users to view tags assigned to any article.
Preconditions	
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to view all tags.</li> <li>2. System displays all public tags</li> <li>3. User may select a tag to view the articles assigned with this tag.</li> <li>4. System displays a list of articles with this tag.</li> </ol>
Alternative Courses	<p>User is logged in and chooses to display all tags.</p> <ul style="list-style-type: none"> <li>• System displays tags assigned by the user followed by all public tags.</li> </ul> <p>User is logged in and chooses to display their own tags.</p> <ul style="list-style-type: none"> <li>• System displays tags assigned by the user followed by all public tags.</li> </ul>
Exceptions	
Implementation Notes	
Interface Object/s	<a href="#">Tag Cloud</a> <a href="#">Tagged Articles List</a>

Back to [List of use cases and interface objects](#)

### Use case: View comments

Actors	User
Description	<p>This use case allows users to view comments assigned to a single article or to view comments they themselves have created on any article.</p> <p>Users can not see comments assigned to multiple articles by other users, except for comments visible in use case: View user profile.</p>
Preconditions	User is viewing a single article.
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User chooses to view comments.</li> <li>2. System displays the public comments.</li> <li>3. User selects a comment.</li> <li>4. System shows the comment's history i.e. all versions of the comment.</li> </ol>
Alternative Courses	<p>User is logged in while viewing a single article.</p> <ul style="list-style-type: none"> <li>• System displays the public comments plus any non-public comments the user has created for this article.</li> </ul> <p>User is logged and chooses to view all their own comments.</p> <ul style="list-style-type: none"> <li>• System displays all the comments the user has created for any article..</li> </ul>
Exceptions	
Implementation Notes	
Interface Object/s	<a href="#">Show Comments</a> <a href="#">Your Comments</a>

	<a href="#">Comment History</a>

Back to [List of use cases and interface objects](#)

### Use case: View user profile

Actors	User
Description	This use case allows users to view a user's public input to the system.
Preconditions	
Postconditions	
Basic Course	<ol style="list-style-type: none"> <li>1. User selects a user profile to view.</li> <li>2. System displays the user profile with public tags and comments and the most recent text corrections the user has added.</li> </ol>
Alternative Courses	<p>User is logged in and selects their own user profile to view.</p> <ul style="list-style-type: none"> <li>• System displays the user profile with non-public as well as public tags and comments and the most recent text corrections the user has added.</li> </ul>
Exceptions	
Implementation Notes	
Interface Object/s	<a href="#">Profile for User</a>

Back to [List of use cases and interface objects](#)

### Interface objects and business rules

#### Standard header for the site

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Banner	Contains name of the service: Australian Newspapers
Home link	
About Us link	Links to an About screen with info about the prototype itself and the project.
Browse link	Links to Browse screen
Help	Links to Help screen
Login/Signup	Login/Signup appears when the user is not logged in. When the user is logged in the message 'Logged in as [username]' appears and the Login/Signup link changes to Logout.
Logout	When the user chooses the Logout link the user's session is terminated and the Login/Signup link places Logout.
Search box	The search box is part of the standard header and so appears on every screen. See Search functionality below for text parsing and search processing rules:
Search Results link	This link appears to take a user back to the last Search Results. In fact it uses a cookie set for the last search performed by the user to re-do

	the search and re-present the search results. (so if the data has been updated meanwhile there is a small chance the search results could be different). The cookie expires when the browser is closed.
	Search Results link only appears if the user has performed a search in this browser session.
	Search Results link does not appear on the Search Results page itself.
Advanced Search	Links to Advanced Search

Back to [List of use cases and interface objects](#)

## Signup

<i>Element</i>	<i>Comments, business rules etc</i>
Choose your user name	mandatory; between 4 and 32 characters long; it contains only alphabetic characters, digits, dashes and full stops; it starts with an alphabetic character; it hasn't been used by anyone else
Password	mandatory; between 4 and 20 characters long
Repeat password	mandatory; must be the same as the password
Email address	mandatory
Are you a human?	mandatory; captcha challenge. If user can't read it, they can refresh the page to try another image or click on a link to listen to the same characters being read.
Accept the terms of use	mandatory; Links to Terms of Use and privacy policy are provided.

Back to [List of use cases and interface objects](#)

## Login

<i>Element</i>	<i>Comments, business rules etc</i>
Username	NLA staff will use their LAN logon. If duplicated by a public username, NLA staff will need to add staff: in front of their LAN logon e.g. staff:blee Other (public) users will use the username they nominated when they signed up.
Password	NLA staff will use their LAN password. Other users will use the password they nominated when they signed up.
Login button	Submits username and password to the system for authentication.
Forgotten password link	Links to <a href="#">Reset Forgotten Password</a> screen.
Terms of Use link	Links to Terms of Use
Privacy policy link	Links to privacy policy on the NLA website
Signup link	Links to <a href="#">Signup</a>
Explanatory text	
Disclaimer	

Back to [List of use cases and interface objects](#)

## Home

<i>Element on screen</i>	<i>Comments, business rules etc</i>

Standard header	Without search box
Featured article	"On this day" – an automatically selected article which changes when the screen is refreshed.
	Rules for selection? -favours news articles with illustrations.
	Presentation of page image as for <a href="#">Article View</a> screen.
	Citation links to the <a href="#">Page View</a> of the first page of the issue.
Read this article link	Links to the <a href="#">Article View</a>
Find an Article	Search box. See <a href="#">Search functionality</a> below for text parsing and search processing rules:
Search Results link	This link appears to take a user back to the last Search Results. In fact it uses a cookie set for the last search performed by the user to re-do the search and re-present the search results. (so if the data has been updated meanwhile there is a small chance the search results could be different). The cookie expires when the browser is closed.
Advanced Search link	Links to <a href="#">Advanced Search</a>
By Title	Each title links to the <a href="#">Newspaper View</a> for that title.
Show all titles	Links to <a href="#">Newspaper Titles List</a>
By State	Regions on the map and the text links to <a href="#">Newspaper Titles List</a> for the selected state.
By Date	Allows users to select a year, then month, then day. Only dates with newspaper issues available can be selected.
User Login	Allows user to login by entering their username and password.
Recently added comments	A list of the most recently added comments and the user who added them.
	Each note links to the <a href="#">Article View</a> for the article containing the note.
Your Comments link	Appears if user is logged in. Links to <a href="#">Your Comments</a> .
Recently added tags	A list of the most recently added tags. Each tag links to the <a href="#">Tagged Articles List</a> screen.
Link to all tags	Links to <a href="#">Tag Cloud</a> .
Your Tags link	Appears if user is logged in. Links to <a href="#">Tag Cloud</a> .
Top text correctors	Acknowledges users who have done the most text corrections. Shows username and the number of corrections done. Username links to <a href="#">Profile for User</a> . Corrections by anonymous users should not appear here.

Back to [List of use cases and interface objects](#)

## Search functionality

Feature	Comments/examples	Sample results
Search is case insensitive	except for upper case AND, OR or NOT which are treated as Boolean operators.	
Multiple word search	<code>climate change</code> finds articles with the words 'climate' and 'change'	2148 matches <code>climate AND change</code> gives same result.

Automatic stemming	Some automatic stemming is applied to word and phrase searches eg climate also searches as climat*.	
Phrase search	"climate change" finds the phrase 'climate change'	19 matches. Finds eg 'climate, change', 'climatic changes', 'changeable climate' in addition to 'climate change'.
AND operator	climate AND change finds articles with the words 'climate' and 'change'	2148 matches
OR operator	climate OR change finds articles with either the word 'climate' or the word 'change' or both. 'OR' must be in upper case.	37306 matches (2064 matches with 'or' in lower case)
NOT operator	climate NOT change finds articles with the word 'climate' where the word 'change' does not occur.	climate 8473 matches climate NOT change and also climate - change and also climate - change 6325 matches Note: climate-change gives same results as "climate change".
Parentheses	To change order of precedence in Boolean expressions eg "Latin America" AND (elections OR politics) NOT (BRAZIL OR MEXICO)	
Proximity	pottery and glass within 7 words of each other: "pottery glass"~7	
Truncation and wild card	climat* finds words beginning with 'climat' euthana*a finds words beginning with 'euthan' and ending in 'a'	climat* 8670 matches climate 8473 matches climat* change 2199 matches
Search contains something resembling a date	Dates can be single or range, in various formats 12 sep 1945 12/09/1945 12/9/1945 12 september 1945 Or just months and years june 1945 dec 1888 Or just years 1888  And separated with space or dash for a range june 1945 - aug 1945 01/01/1900 - 17/8/1912	war france 12 oct 1916 14 march 1917 hitler argus may 1945 germany april 1913 - nov 1916 germany april 1913 - nov 1916 hobart mercury bomb 3/9/1945 jacques 1858 sa advertiser chicken soup 1913-1928 war the bloody tizer
Search contains	We also try to guess titles based on title variations now stored and updatable with the QA title,	

something resembling a newspaper title	<p>"The Argus" is matched by "The Argus" and "Argus"  "The South Australian Advertiser" is matched by "The South Australian Advertiser" or "South Australian Advertiser" or "SA advertiser" or "advertiser" or " the bloody tizer"  matching is based on a regular expression like this:</p> <pre>(?:\b(?:the south australian advertiser) (?:south australian advertiser) (?:sa advertiser) (?:advertiser) (?:the bloody tizer)\b)</pre>	
Fielded searches	<p>Some fielded searching is possible eg</p> <p>NOT category:Advertising will exclude advertisements from the search results;</p> <p>yyyy:1880 finds articles published in 1880;</p> <p>yyyymmdd:18800209 finds articles published on 9 Feb 1880;</p> <p>firstpageseq:4 finds articles which occur on, or which start on, page 4 eg court reports (firstpageseq:4 OR firstpageseq:5)</p> <p>headingsauthors:climate finds articles with 'climate' in the title, subtitles or author.</p> <p>fulltext:climate finds articles with 'climate' in the title, subtitles, author or text.</p> <p>batchid:67 finds articles from batch 67 (in March 2008 implemented on base only, not on ndpprototype).</p>	
Relevance ranking	<p>Search results are displayed in order of relevance. This currently works on the following principles:</p>	
	<p>Inverse document frequency i.e. rarer words are weighted more than common words.</p>	<p>eg a document with 17 occurrences of 'on' and 1 occurrence of 'smartboard' ranks lower than one with 17 occurrences of 'smartboard' and 1 occurrence of 'on'.</p>
	<p>Headings (title/subtitle/author) are weighted more than full-text.</p>	
	<p>Length of article and density of words</p>	<p>eg a short article with 3 occurrences of a word ranks higher than a longer article with 3 occurrences of the word.</p>
	<p>Words in close proximity are weighted.</p>	<p>eg a phrase (two words together) ranks higher than where two words occur within 5</p>

		words of each other, which ranks higher than where two words occur within 20 words of each other.
	Articles are ranked lower if they are Advertising (10% as important) or Detailed lists (20% as important). For efficiency this has been done on ingest. If we want to change this the whole database has to be re-indexed.	eg if three articles had each scored 500 to this point, a news article's final score would be 500 but an advertisement's final score would only be 50 and the detailed list's final score would be 100.

Back to [List of use cases and interface objects](#)

## Browse

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Browse Issues	This section of the screen allows users to start browsing by newspaper title or to browse by date.
By title	Shows 5 titles to select from with an option to show all titles. Individual titles link to <a href="#">Newspaper View</a> . 'Show all titles' links to <a href="#">Newspaper Titles List</a>
By state	Each state links to <a href="#">Newspaper Titles List</a>
By date	Users select a year, then a month, then a day.
Year selection	Users may only select a year from the range of years for which digitised issues are available (all newspapers)
	Users may only select one year.
	Selecting a year populates the month selection element.
Month selection	Users may only select from the months for which digitised issues are available (all newspapers)
	Users may only select one month.
	Selecting a month populates the issue selection element (all newspapers).
Issue selection	Users may only select from the digitised issues available. (An issue of a newspaper is defined by its date of publication).
	Users may only select one issue.
	Selecting an issue takes users to <a href="#">Page View</a> of the first page of the issue.
Browse articles	This section of the screen allows users to browse articles by selecting a category or a tag.
By category	Lists all categories available (Advertising ; Detailed lists, results guides ; Family Notices ; News). Selecting a category executes a search for all articles in that category.
By tag	Lists the most common tags and an option to show all tags. Each tag links to Tagged Articles List. 'Show all tags' links to <a href="#">Tag Cloud</a> .

Back to [List of use cases and interface objects](#)

## Newspaper View

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	Note: Search box searches all articles in beta, not restricted to this newspaper.
By year	Graphical representation of number of issues and range of years which have been digitised for this newspaper. If you click on part of the graph it changes the year selected in the 'By date' section below.
By date	This is like the Browse screen, with years, months and issues restricted to this newspaper.
Cite this title	Contains NLA persistent identifier for this newspaper e.g. <a href="http://nla.gov.au/nla.news-title13">http://nla.gov.au/nla.news-title13</a>
Newspaper information	Contains covering text about the newspaper e.g. acknowledgements. Displays first four lines of Wikipedia entry (dynamically imported) with a link to the full entry in Wikipedia. The link may be edited in the NDP Content Management System (Edit title).
Bibliographic record	Dynamically imports MARCrecord from National Library catalogue (Voyager). (Libraries Australia doesn't have a service we can use to do this)
	Displays appropriate parts of the MARC record in a user-friendly way.
Links to library holdings	Links to Libraries Australia detailed display, which includes holdings information, of records for the print, microform and electronic copies.
<i>Implementation notes</i>	<p>Title fragments for title/title-browse page may be edited by NDP staff with access to the appropriate directory.</p> <p>There is a file for each title; the filename is the number in the title's persistent identifier e.g. 13.txt for the Argus.</p> <p>This file is a template - everything you enter will get copied to the page except for a string of @@@&lt;something&gt;@@@ Which is assumed to wrap the voyager number - eg @@@1564333@@@</p> <p>The voyager number is extracted, used to query voyager and the contents after formatting used to replace the @@@...@@@</p> <p>-----</p> <p>The first 3 paragraphs or 200 characters of a wikipedia article can be shown by inserting something like this in the title templates:</p> <pre>&lt;DIV&gt; &lt;B&gt;From Wikipedia:&lt;/B&gt; &lt;DIV style='margin:1em;font-size:smaller'&gt; \$\$\$SHOWWIKI=The_Argus_%28Australia%29\$\$\$ &lt;a href='http://en.wikipedia.org/wiki/The_Argus_%28Australia%29'&gt;&lt;i&gt;Mor e...&lt;/i&gt;&lt;/a&gt; &lt;/DIV&gt; &lt;/DIV&gt;</pre> <p>The important bit is the \$\$\$SHOWWIKI= and \$\$\$ markers which</p>

	<p>enclose the page name on wikipedia (the bit after the <a href="http://en.wikipedia.org/wiki">http://en.wikipedia.org/wiki</a> in the wikipedia URL)</p> <p>This marker text will be replaced with the first 1 or 2 3 paragraphs of text from wikipedia - the number of paras is increased until the character limit of 200 is reached (easily changed).</p> <p>The extracted text is cached, so changes in wikipedia wont be reflected in the NDP until the cache is cleared (currently when the application is restarted, or a big garbage collection occurs).</p> <p>-----</p> <p>Michelle Murray from the newspaper reading room suggested the following fields from the catalogue record for display.</p> <p>Title: 245 XX abhcfknps, 591 XX a</p> <p>Publisher: 260 XX abc</p> <p>Description:300 XX abcefg 507 XX ab 753 XX abc 755 XX axyz</p> <p>Notes: 500 XX a 501 XX a 502 XX a 504 XX a 513 XX ab 516 XX a 518 XX a 521 XX a 522 XX a 524 XX a 525 XX a 530 XX abc3 533 XX abcdefmn 535 XX abcdg 544 XX abcden 546 XX a 547 XX a 550 XX a 556 XX az 562 XX abcde 565 XX abcde 567 XX a 580 XX a 581 XX a 583 XX az</p> <p>Life Dates:362 XX a</p> <p>Later Title:785 XX abdstxz</p> <p>Former Title:247 XX abfghnpx 780 XX abdstx</p> <p>Also Titled:246 XX abnpfgh 730 XX adfghklmnoprstx 740 XX anp</p> <p>Place:752 XX abcd</p>
--	---

Back to [List of use cases and interface objects](#)

## Advanced Search

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Keywords	all of these words
	with the exact phrase
	with any of these words
	without these words
(where)	anywhere in article
	headings only (titles and subtitles)
	headings, author or 1 <sup>st</sup> 4 lines (i.e. the text rekeyed/cleaned by Apex - 1 <sup>st</sup> 4 lines are cleaned in news articles only)
	captions only (only applies to illustrations)
Publication date	from Day / Month / Year
	to Day / Month / Year
Newspaper Title & Location	Users can select all newspapers in a state, individual titles or a mixture.
Article category	News ; Advertising ; Family Notices ; Detailed Lists, Results, Guides
	Superseded categories of Display Advertising and Classified Advertising

	have been amalgamated into Advertising for searching (but are still searchable by title).
Article length	All ; < 100 words ; 100-1000 words ; > 1000 words
Illustrated	Restrict to illustrated articles
Sort	Option to sort by relevance (the default) or date.(earliest first ; latest first).
Batch(es)	Single batch number or range e.g. 80-85. This option is only visible to NDP staff (i.e.NLA staff with the appropriate role/s in the user management system.)
Search button	Executes the search and presents <a href="#">Search Results</a> .

Back to [List of use cases and interface objects](#)

## Search Results

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	Does not include the Search Results link.
	Search box contains option to search within these results or do a new search (the default is new search).
Results message	Shows terms used in search.
	Includes the number of matches found.
	Shows any filters applied to the search with an option to remove each filter, which re-executes the search without the filter.
Results list	Shows 10 results per screen
	Although not a feature for users, it is possible to change the number of results per screen eg to change to 100, add <code>&amp;maxDisplay=100</code> to the URL. It's also possible to specify where to start from eg to start from result 400, add <code>&amp;startFrom=400</code> to the URL.
	Users can jump to previous/next set of 10 results
	Users can jump back to the first page of results.
	Users can jump to last page of results.
	Users can navigate to any other page of results.
	Each result is numbered.
	Each result is an article citation
	Articles from issues still being processed or QA'd are highlighted in pink. These articles are only available to authorised users to view.
Refine search	Goes to <a href="#">Advanced Search</a> populated with user's search terms.
Sort options	By relevance (the default); by date earliest first; by date latest first.
Article citation	Title of article. Links to <a href="#">Article View</a> screen.
	Subtitle/s of article
	Author/s of article
	Source of article: newspaper title date (day of week, day of month, month, year)

	page number
	article category
	length of article
	<p>Snippets showing search terms in context:</p> <p>Non-preferred words are from snippets: the presence of an APEX correction "trumps" an OCR word and the presence of a human-non-APEX correction "trumps" both.</p> <p>Reconstructed hyphenated words are ignored.</p> <p>If something appears a bit odd it may be due to word misalignment during corrections (word boundaries may be different in the original OCR).</p>
	Users can see if the article has been corrected
	Users can see the most recent tags added.
Clusters / Facets	<p>Shows the distribution of articles from the result set in the following clusters:</p> <p>Newspaper Tilt  Category  Illustrated  Decade  Word Count</p>
	Users can narrow their results to a particular cluster.
Results from and links to other services	<p>NBD (Library Labs version)  PictureAustralia  ...</p>

Back to [List of use cases and interface objects](#)

## Article View

e.g. <http://ndpbeta.nla.gov.au/ndp/del/article/1040368>

Note re image:

Oxide Interactive recommended the article view consist of rearranged snippets, as in the Save /Print Article options. Reasons for retaining the article in the context of the page are [from Kent]:

- it preserves the newspaper layout . This isnt just "nice"/"faithful"/"accurate", showing the article as its creator intended and as it was originally viewed, it is sometimes contextually useful ("shown in the photo to the left" isnt rendered meaningless by us changing the original layout).

- serendipity - Im constantly surprised and delighted by the content I accidentally come across (out of the corner of my eye) when viewing something else (the thing I came to see)

- an article snippet view makes it harder for users to recover from inevitable article zoning errors, where an article is truncated by APEX. Seeing the article in the page view makes these errors less damaging

Secondary considerations were: If snippets were presented, derivatives would have to be generated on ingest and stored for every article, as dynamic generation (as happens now with

Save / Print Article JPGs and PDFs) would compromise the speed of delivery. Also if pre-generated:

We would need to change ingest workflow (can be done)

Service would be slow with pre-generated PDFs so would use JPEGs.

With JPEGs we would need to decide if users get to see a single size and therefore no zooming or if more than one size so they can zoom we would need to know sizes for tiling - takes extra storage (but we have space).

-----

The article, page and issue views underwent major change between ndpprototype and beta. The purpose of these changes was to better integrate the article, page and issue views and improve navigation between articles, pages and issues.

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Article citation, i.e. Article title, subtitle/s and authors, abstract, article category, article length	Were in ndpprototype but not in beta or the current system. This saved space and eliminated the citation data appearing twice (the Apex rekeyed data and the same data in the OCR text).  The rekeyed data still appears for each article in the Page View.
Newspaper title	With an 'about' link to <a href="#">Newspaper View</a>
Date of issue	In beta the date is not a link. (In ndpprototype the date linked to an issue view, which doesn't exist in beta.)
Links to next issue and previous issue	Links to <a href="#">Page View</a> for page 1 of the next or previous issue
Page number	Shows number of the page currently being viewed.
	User can select any other page in the same issue to go to.
Link to next page and previous page	Links to Page View for the next page or previous page. Next / previous page arrows do not appear if at last / first page.
View entire page	Links to Page View for the current page.
Page image	Page image is loaded at default zoom level (35%).
	Article is white, rest of page is greyed out though still readable.
	Article is placed at top left side of viewer.
	Users can navigate to next and previous zones of the article, including where an article continues on another page (problem in IE but works in other browsers)
	Users can pan around the page image. Users can click and drag directly on the image. Users can use the scroll bars. In ndpprototype users could also move an active area indicator on a thumbnail of the image. This was removed in beta and replaced with scroll bars. The arrow keys also pan the image (slowly and only in Firefox at present).
	Users can zoom by choosing from eight zoom levels (5, 15, 25, 35, 45, 55, 65 and 100 %) or choosing to go up or down one zoom level at a time. (In Firefox only you can also use the + and - keys on the number pad.) Double-clicking centres the article and zooms in (+) one level.

	<p>In ndpprototype: When zooming the page image was anchored at the top left corner (as in Acrobat reader for PDF documents) (not at the centre as in Google maps).</p> <p>This was changed to anchoring at the centre on Oxide's recommendation.</p>
	<p>In ndpprototype: Users could alter the width of the image viewer by clicking and dragging the left hand side of the viewer (frame splitter). Also this is the way to hide the OCR text.</p> <p>In beta there is no way to do this because the frame splitter won't work with the Oxide styles (problem with divs and width of background images).</p>
	<p>Users can get help on how to use the page viewer (pan, zoom etc). (In beta this is now in the general Help and also on the Home page near the page image there.)</p>
	<p>When users click on a line of text, that line in the page image is moved to the top left of the image viewer (how well it works is subject to the accuracy of the text coordinates).</p>
Electronically translated text	<p>The initial version is the raw OCR text generated by the OCR software, but the display will show the latest version including subsequent automatic or manual corrections.</p>
	<p>In ndpprototype users could 'synchronise' a zone of the OCR text with its position on the page image. They can click on an icon in the text and the corresponding part of the image will move to the top left corner of the viewer.</p> <p>This was not very intuitive and the icons have been replaced in beta by the ability to simply click on a line of text.</p>
	<p>At present there is no option to hide the OCR text (in ndpprototype the user could hide the text by moving the a frame splitter lon the left side of the viewer so the viewer took up the whole screen.)</p>
	<p>For articles which span multiple pages, all the OCR text is shown, with an indication of which page the text appears on. The page number is a link and allows the user to bring up the image of that page.</p>
View comments	<p>In ndpprototype comments (notes) were associated with a particular line of text. This was changed in beta so that notes belong to an entire article (because Oxide user testing found users had difficulties understanding comments and placing them.)</p>
	<p>Users can elect to show comments or hide comments.</p>
	<p>Users can then select a comment to view in full. Links to <a href="#">Show Comments</a>.</p>
	<p>Users can see notes they themselves have created, and can also see public notes other users have created.</p>
	<p>There is an Edit link to any note the user can see which links to <a href="#">Add/Edit Comment</a></p>
Add new comment	<p>Users can add a new comment to associate with the article Links to <a href="#">Add/Edit Comment</a></p>
Why may this text have mistakes?	<p>Pops up a window with explanatory text.</p>
Help fix this text	<p>The top link pops up a window with explanatory text. The lower links which appear when you hover over a line in the OCR</p>

	text invokes the Correct the Text use case.
Correct the text	Users can choose a line of text to correct (by clicking on 'Help fix this text' link which appears when they hover over a line.) or users can choose the 'Help fix this text' link at the top of the text. (Correction is line by line so word positions in the page image are preserved and can be correlated with the old text, and highlighting of search terms can still occur.)
	<p>A full screen edit mode is presented which is like the Article View screen but has header, citation, and print and save buttons removed. Each line of text appears in an editable text box and edit control buttons appear, "Exit", "Cancel" and "Insert Symbol".</p> <p>"Undo Line" and "Save" are greyed out until the user has made a change to the text.</p> <p>"Exit" changes to "Save and Exit" when a user has made a change to the text</p> <p>The first line of text is focused if the top "Fix this Text" button was clicked, otherwise the line the user clicked is focused.</p> <p>"Undo Line" undoes all the changes that have been made to the current line.</p> <p>"Insert Symbol" presents a menu of commonly used symbols (e.g. the pound symbol) and allows users to select one to insert in the text.</p> <p>"Save" saves the changes and allows the user to continue editing.</p> <p>"Save and Exit" saves the changes, closes the edit screen and returns to the article view screen.</p> <p>"Cancel" closes the edit screen and returns to the article view screen without saving changes made since the last save.</p> <p>Colour coding provides an indication of which lines have been changed and not saved, and which lines have been changed and saved.</p> <p>A link to Help and information about Keyboard shortcuts is provided on the edit mode screen as well as the article view screen, and the Help has been updated and reformatted to make it easier to follow.</p>
	The selected line moves to the top left corner in the image viewer.
	Users can type replacement text for the line (in a text box).
	Multiple spaces in the text box are removed.
Adding missing text	<p>Functionality of NDPbeta production system is retained for anonymous users, i.e. they are able to add text to the preceding line. They are limited to adding 10% of the current length of the line with each Save transaction. An "End of line reached" message is presented when the 10% limit is reached.</p> <p>Logged in users are also presented with an "End of line reached" message when they reach the 10% limit but they are also prompted</p>

	<p>"Move to the next line or click to add missing text". This facility is restricted to logged in users to mitigate the risk of vandalism. Note that the 10% limit is still required for logged in users because this serves as a reminder that they should go to the next line when correcting text normally. Clicking the link will allow them to add up to another 400 characters before the "End of line" message is re-presented. An absolute limit of 1200 characters can be added to a line within a Save transaction. Once this amount of characters have been added to a line, subsequent edits are limited to 10%.</p> <p>Users are advised (in help text) to add missing text to the end of the preceding line. If they add text to the beginning of the next line, it will push all the word co-ordinates out.</p> <p>So that users can edit without scrolling, if a line is long enough that the text needs to wrap, the text area will resize to accommodate the text. Carriage return has been disabled in the textarea because of the use of Enter as a keyboard shortcut to move to the next editable line.</p> <p>As textareas do not support a HTML maxlength, the maximum line length is checked with javascript - when typing up to the limit the text will trim to the max length and a 'maximum length' message appears.</p> <p>A solution for adding text to articles with 0 words has not been implemented.</p>
	<p>Keyboard shortcuts:</p> <p>The principle to use browser defaults where possible for keyboard behaviour was abandoned because different browsers behaved differently and it was thought better to have consistent behaviour across browsers. It was thought acceptable to depart from default browser behaviour because OCR text correction is specialised functionality unlike general web browsing. However contradictory behaviour for a keyboard shortcut was to be avoided.</p> <p>The keyboard behaviour implemented is:</p> <p style="padding-left: 40px;">TAB allows you to move to the next line while Shift-TAB allows you to move to the previous line.</p> <p style="padding-left: 40px;">Enter allows you to move to the next line.</p> <p style="padding-left: 40px;">Up and Down arrow keys also allow you to move from line to line.</p> <p style="padding-left: 40px;">Left and Right arrows allow you to move within a line or a text area. Use the Home and End keys to position the cursor at the beginning and end of the input field respectively.</p> <p style="padding-left: 40px;">Windows: Alt + s focuses the Save button, Alt + e focuses the Exit/Save &amp; Exit button and Alt + x focuses the Cancel button. Hit Enter to perform the action. (These rely on browser "access key" functionality and may not work with all browsers.)</p> <p style="padding-left: 40px;">MacOS X: Ctrl + s is equivalent to Save, Ctrl + e to Exit/Save &amp; Exit and Ctrl + x to Cancel.</p>

	<p>Using Enter or Down arrow on the last input does not cycle to the top input but retains focus in the last input. TAB does not retain focus in the last input as this would be contradictory behaviour compared with the TAB key's browser default behaviour.</p>
	<p>System escapes &lt; and &gt; characters in OCR corrections (i.e. change &lt; to &amp;lt; and &gt; to &amp;gt;) because it is assumed it is never valid to have HTML in OCR corrections and to avoid mischievous insertion of code.</p>
	<p>Length limit of OCR corrections is 200,000 characters at one time. (Users can save their corrections then do some more. They should be encouraged to save their work at regular intervals in any case.)</p>
<p>Other features related to text correction</p>	<p>If when saving an error occurs on the server an alert will display the error message. The line will be styled with a reddish background so users can identify which line failed to save</p> <p>The number of corrections made to the article has been added to the corrections information at the top of the article text.</p> <p>A check is done for unsaved changes upon leaving the page. The user is notified that unsaved changes exist and the message allows them to remain on the page to save them. This uses browser functionality and the browser default message has been modified to provide specific information.</p> <p>To indicate the end of the page, after the last input a heading has been added: "End of Electronically Translated Text" or "Continue to Page X" for multiple page articles.</p> <p><i>Abuse of text correction facility</i></p> <p>Abuse has not been a problem to date. However the possibility of abuse has been, and will continue to be, kept in mind. Strategies lean toward containing the damage rather than preventing abuse altogether.</p> <p>Text correction will continue to be unmoderated as previously. The need for moderation will continue to be evaluated. Moderation by the community is the preferred model and an online facility to support this may be provided in future if required.</p> <p>If abuse is reported, it is possible for system administrators to roll back transactions for a specified article, user, and/or time.</p> <p>Reports may be run on the database to detect certain types of abuse if required e.g. by looking for certain words or line lengths.</p> <p><i>Some technical details</i></p> <p>The line length limit will be implemented in the javascript.</p> <p>Previously the save action of a line was completely synchronous i.e. once users clicked save, they couldn't continue editing until the server responded to the save request. The save action has been enabled to</p>

	<p>use an asynchronous request that allows the user to continue on while the actual saving happens in the background. As such, if the external user is experiencing network congestion or the server load is high, any delays in committing the line changes will not disrupt the user from continuing to edit further lines.</p> <p>Response time issue on loading the edit page for a long article was addressed.</p>
Show corrections	'Show corrections' links to <a href="#">Show Corrections</a> .
	Every time changes are saved, a new version of the full text of the article is kept in the database in case there is ever a need to 'roll back' to an earlier version (eg due to vandalism). However there is no facility at present for users to view the full text of earlier versions of the article – users can only view the lines which have been corrected.
Revise corrections to the text	Users can revise corrections made by any user simply by choosing that line of text and making further changes.
View tags	Users can see what tags have been applied to this article. They can only see tags which are public or which they themselves have created.
	Each tag is a link to <a href="#">Tagged Articles</a> List
Add / edit tags.	Users can choose to add or edit tags Links to <a href="#">Edit Tags</a> .
Link to view all changes made by the last person to correct the article text.	Links to <a href="#">Profile for User</a> .
Save as Image	Links to <a href="#">Save Article as Image</a>
Print	Links to <a href="#">Save Article as PDF</a> and also brings up a print dialog box.
Save as PDF	Links to <a href="#">Save Article as PDF</a> without the print dialog box.
Link to page image in QA system	Was in ndpprototype. Not in beta at present. Was Only visible to authenticated users with role of ndpqa.
Cite this article	<p>Cites the persistent identifier (PI) for this article.</p> <p>Users who use the browser's bookmark function to save the url of the Article View will be able to retrieve this article again unless the search &amp; delivery system changes. However the url is not guaranteed to be persistent in the long term, unlike the PI.</p>

Back to [List of use cases and interface objects](#)

## Page View

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Newspaper title	With an 'about' link to <a href="#">Newspaper View</a>
Date of issue	In beta the date is not a link. (In ndpprototype the date linked to an issue view, which doesn't exist in beta.)

Links to next issue and previous issue	Links to <a href="#">Page View</a> for page 1 of the next or previous issue
Page number	Shows number of the page currently being viewed.
	User can select any other page in the same issue to go to.
Link to next page and previous page	Links to <a href="#">Page View</a> for the next page or previous page. Next / previous page arrows do not appear if at last / first page.
Link to page image in QA system	Only visible to authenticated users with role of ndpqa.
Page number	
List of articles	Articles are presented in 'reading order' according to where their first zone is placed. Starts from top left hand corner of the page, flows down to the bottom of the column, then across to the next column.
Article title	Links to <a href="#">Article View</a> . (Article titles are rekeyed by Apex for categories News and Detailed results, lists, guides. Otherwise Apex assigns the article category as the article title).
Article subtitle	Article subtitle as defined and rekeyed by Apex.
Article abstract	This is the first four lines of an article where the OCR text has been rekeyed or 'cleaned' by a human operator. (does not apply to Advertising, Family notices, Detailed lists and guides).
Article category	Article category assigned by Apex.
Article length	
Article continues on another page or is continued from another page	Users can see if an article continues on another page or is continued from another page.
	A link to the <a href="#">Page View</a> of the next/previous page the article appears on is provided.
Sync article to page image	When the icon is clicked, repositions the page viewer to the article ( places the article at top left side of viewer).
Page image	Page image is loaded at default zoom level (15%).
	Top left corner of page is placed at top left corner of viewer.
	Users can pan around the page image. Users can click and drag directly on the image. Users can use the scroll bars. The arrow keys also pan the image.
	Users can zoom by choosing from eight zoom levels (proposed: 5, 15, 25, 35, 45, 55, 65 and 100 %) or choosing to go up or down one zoom level at a time. (In Firefox only you can also use the + and – keys on the number pad.) Double-clicking centres the article and zooms in (+) one level.
	When zooming the page image is anchored at the centre on Oxide's recommendation.
	Users can alter the width of the image viewer by clicking and dragging the left hand side of the viewer (frame splitter).
	Users can get help on how to use the page viewer (pan, zoom etc). (In beta this is now in the general Help and also on the Home page near the page image there.)

	When navigating to the previous or next page, the page image is displayed at the user's currently selected zoom level.
	When user hovers over an article on the page, the boundaries of the article are highlighted.
Save as PDF	Links to <a href="#">Save Page as PDF</a> 'Save as PDF' just opens the PDF – we can't open a print dialog box for PDFs, so there is no point putting 'Print' in the label. Also pages won't be readable at A4 (A4 is the usual default paper size people have their printers set at). The PDF pages were sized for A3 because A4 is useless for reading a page. Using the Adobe reader users can resize the PDF for printing and can specify a part of the page to print.
Download issue as PDF	Allows users to download all the pages in the issue as a PDF document.
Cite this page	Cites the persistent identifier (PI) for this page.  Users who use the browser's bookmark function to save the url of the Page View will be able to retrieve this page again unless the search & delivery system changes. However the url is not guaranteed to be persistent in the long term, unlike the PI.

Back to [List of use cases and interface objects](#)

## Issue View

This is now the same as the [Page View](#) of the first page of the selected issue. There is no separate 'issue view'.

There is an option on the Page View screen to **download an Issue as a PDF**

Back to [List of use cases and interface objects](#)

## Save Page as PDF

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	There is no standard header
General	Opens a PDF version of the page. This PDF is created on ingest of the files after OCR processing and is designed to print the complete page on an A3 sheet of paper without adjustment.
Bookmarks	Bookmarks are displayed for each article on the page.
Print page	Users can use their browser's print function to print a copy of the page.
Save / download page	Users can use their browser ('Save As' function in IE) to save / download a copy of the PDF file.
	Many functions are possible with recent versions of Adobe reader, some of the most useful of which are:
	Search for text in the page.
	Print the whole page and shrink or enlarge it to fit on any size paper.
	Select and print a part of the page and shrink or enlarge it to fit on

	any size paper.
	Select and enlarge any part of the page for easier reading on the screen using various methods (from the Tools menu).
Footer	Footer includes identification of the National Library of Australia as the source and the Persistent Identifier.

Back to [List of use cases and interface objects](#)

## Print Article

e.g. <http://ndpbeta.nla.gov.au/ndp/del/printArticleJpg/1040368/3?print=y>

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	There is no standard header
General	As for <a href="#">Save Article as Image</a> below, but also pops up a Print dialog box.

Back to [List of use cases and interface objects](#)

## Save Article as Image

e.g. <http://ndpbeta.nla.gov.au/ndp/del/printArticleJpg/1040368/4?print=n>

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	There is no standard header
General	Opens a new window containing some covering text and a JPEG image
JPEG image	The JPEG image is dynamically generated at the currently selected zoom level.
	The zones / columns are laid out to fit on as few A4 pages as feasible while maintaining reading order.
	The JPEG image contains all zones in the article including where an article spans multiple pages in an issue.
Covering text	Explains how many A4 pages the image will print on and that the user can select another zoom level if the print is too small (e.g. if too small to read) or too large (e.g. if takes too many pages to print).
Print article	User can use their browser's print function to print the image. (Note that if the user has set their own margins for printing, the full width of the article may not fit on a page.)
	The covering text is not printed.
Citation	The newspaper title, issue date, and page number/s are cited below the covering text but, unlike the covering text, appear when the image is printed.  The newspaper title, issue date, and page number/s are also cited in the <title> element of the html page. The <title> element will usually appear at the top of the copy printed from the browser.
Save / download	Users can use their browser to save / download the image (in IE it is

image	the 'download picture' function available by right clicking on the image).
-------	--

Back to [List of use cases and interface objects](#)

## Save Article as PDF

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	There is no standard header
General	Opens a new window containing a PDF file
PDF file	The PDF image is dynamically generated at the currently selected zoom level.
	The zones / columns are laid out to fit on as few A4 pages as feasible while maintaining reading order.
	The PDF file contains all zones in the article including where an article spans multiple pages in an issue.
	The PDF file is automatically divided into pages. Each page should print without adjustment on A4 paper.
Print article	Users can use their browser's print function to print the PDF file.
Save / download article	Users can use their browser to save / download a copy of the PDF file ('Save As' function in IE).
	Many functions are possible with recent versions of Adobe reader, some of the most useful of which are:
	Search for text in the article.
	Print the whole article and shrink or enlarge it to fit on any size paper.
	Select and print a part of the article and shrink or enlarge it to fit on any size paper.
	Select and enlarge any part of the article for easier reading on the screen using various methods (from the Tools menu).

Back to [List of use cases and interface objects](#)

## Show Comments

<i>Element on screen</i>	<i>Comments, business rules etc</i>
General	Part of Article View screen which expands when Show Comments link is selected. Users can only see public comments or comments they themselves have created.
Comment details	The latest version of each comment is shown. The version number, the user who created the note, and date and time of creation are displayed with the comment.
View comment history	The comment is hyperlinked to <a href="#">Comment History</a> .

Edit	Users can revise a note whether it was originally written by them or another user. Links to <a href="#">Add / Edit Comment</a>
------	--

Back to [List of use cases and interface objects](#)

## Add / Edit Comment

<i>Element on screen</i>	<i>Comments, business rules etc</i>
General	This screen is a popup window in Article View.
Editing (adding or revising) a comment	Users can use some html markup to format their comment.
Rules	Only the following markup may be used: <a href..> <b> <i> <ol> <ul> <li>   <p> All other HTML tags are removed, and then all attributes from tags are removed except HREF from the A (link) tag
	The length limit has been set at 4000 characters per comment.
Make this comment public	Check box, defaults to checked. It means anyone can see the comment. If the user unchecks the box, the comment can only be seen by the user that created the comment. (If the user is not logged in, unchecking the box has no effect; all anonymous comments are public.)
OK button	Users can save the new or revised comment.
Cancel button	Users can cancel adding or revising a comment.
Viewing newly saved notes or revisions	After saving, users receive a confirmation message and can view their new comments or revisions in Article View by using their browser's 'refresh page' function.

Back to [List of use cases and interface objects](#)

## Edit Tags

<i>Element on screen</i>	<i>Comments, business rules etc</i>	
General	This screen is a popup window in Article View.	
Editing (adding or revising) tags	Users can add one or more tags to an article, whether there are already tags or not (tags are separated with a comma).	
	Users can edit tags but only tags they themselves created.	
	Users can delete tags but only tags they themselves created.	
Make these tags public	Check box, defaults to checked. It means anyone can see the tags. If the user unchecks the box, the tags can only be seen by the user that created them. (If the user is not logged in, unchecking the box has no effect; all anonymous tags are public.)	
OK	Users can save their changes..	

Cancel	Users can cancel without saving changes and return to the Article View.	
Rules	Non-word characters are removed from tags in an effort to prevent malicious content.	
	Length limit has been set at up to 30 characters per tag.	
	There is a limit of 50 tags per user per article. (There is no total limit per user or per article).	
Viewing newly saved tags or changes	After saving, users receive a confirmation message and can view their new tags or revisions in Article View by using their browser's 'refresh page' function.	

Back to [List of use cases and interface objects](#)

## Tagged Articles List

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Tag name	
Number of articles	Number of articles with this tag.
Link to all tags	Links to <a href="#">Tag Cloud</a>
List of articles	List of all articles with this tag.
Article details	For each article:
	Article title (linked to <a href="#">Article View</a> )
	Subtitles
	All tags applied, each tag is hyperlinked to Tagged Article List for that tag.
	Newspaper title
	Issue date
	Page number/s
	Article category
	Article length
	Abstract (i.e. 1 <sup>st</sup> four lines; only applies to News articles)
	The user is alerted that more articles are available by displaying the number of articles that would be found with a full text search of the same word/s as the tag. Clicking on the number takes users to the search results page.

Back to [List of use cases and interface objects](#)

## Tag Cloud

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
General	Users can only see public tags and non-public tags they themselves have created. The Tag Cloud screen presented is the same whether user chooses 'All Tags' or 'Your Tags' on the Home screen.
'Your tags'	Displays only the tags assigned by this user.
	Displays the user's username.
	Displays the number of times each tag has been assigned by this user.
	Each tag is linked to <a href="#">Tagged Articles List</a> for that tag but only the articles tagged by this user.
All tags	Displays all public tags which have been added to the system (including ones assigned by the user) as a 'tag cloud' i.e. the more times a tag has been assigned the larger it appears.
	Displays the number of times each tag has been assigned. A tag may be assigned multiple times (by different people) to a single article, hence these numbers don't indicate article counts.
	Each tag is linked to <a href="#">Tagged Articles List</a> for that tag.
	<p>The browse tags page, reached by clicking "All tags" on the Home page, previously displayed a tag cloud of all tags added by users. With so many tags now in the system this page was not usable and was taking too long to load. The browse tags page has been redesigned to make it more efficient to browse tags.</p> <p>The page contains:</p> <ul style="list-style-type: none"> <li>The top 30 most popular tags, with the ability to see more (200)</li> <li>The top 30 most recent tags, with the ability to see more (200)</li> <li>An A-Z index of all tags allowing users to browse manageable chunks of tags starting with each letter of the alphabet</li> </ul> <p>The Your Tags page is similar to the All Tags page but only shows the user's own tags.</p>

Back to [List of use cases and interface objects](#)

## Your Comments

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
List of comments	Lists all comments which this user has added (or edited?) with the following details.
Article details	Article title (linked to <a href="#">Article View</a> ), subtitles, newspaper title, issue date, page number/s of article each comment was added to.
Latest version	Number of latest version of comment (which is also the total number of versions)

Link to all versions	Only if there is more than one version of the note. Links to <a href="#">Comment History</a> .
Creator of latest version	Username which created the latest version (which may not be this user).
Date of latest version	Date and time the latest version was created.
Content of note	Displays the latest version of the note in full.

Back to [List of use cases and interface objects](#)

## Comment History

e.g. <http://ndpbeta.nla.gov.au/ndp/del/annotation/213>

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Link to 'your comments'.	Links to <a href="#">Your Comments</a> .
List of versions	All versions of the comment are shown with the following details.
Version number	
Author of version	Username, or 'anonymous' if the user who made the change was not logged in.
Date and time created	
Content	The full contents of the comment.

Back to [List of use cases and interface objects](#)

## Show Corrections

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Article title	Linked to <a href="#">Article View</a>
Article subtitle	
Newspaper title	
Issue date	
Page number/s	
Article category	
List of changes to the text	Each change contains the following information.
Date/time of change	
User who made the change	Username, or 'anonymous' if the user who made the change was not logged in.
Old lines	Lines of text before the change (Changed from ..)

New lines	Lines of text after the change (Changed to ..)
-----------	--

Back to [List of use cases and interface objects](#)

## Profile for User

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
General	Shows Comments, Tags and Corrections submitted by the user
Comments	As for <a href="#">Your Comments</a> .
Tag cloud	As for 'Your Tags' in <a href="#">Tag Cloud</a>
Text corrections	As for <a href="#">Show Corrections</a> but only shows corrections submitted by the user.

Back to [List of use cases and interface objects](#)

## No Match

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	Terms used in the unsuccessful search are reflected in the Search area.
General	This page contains 'static' content which may be updated from time to time by or at the NDP team's request.
Content	Contains <ul style="list-style-type: none"> <li>• a 'no match found' message</li> <li>• some suggestions for trying again</li> <li>• a link to search tips in Help screen.</li> </ul>

Back to [List of use cases and interface objects](#)

## Reset Forgotten Password

(1) Initial form presented

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
User name	Text box to enter user name. If user has forgotten their user name they are reminded they can check their old emails. The email sent on signup contains the user name.
Captcha challenge	Text box to enter characters read in the captcha. Includes a link to an audio version.
Initiate password reset	Button submits the form. If input is valid and user name exists, a message is presented which tells user to check their email and to click on the link or paste the url contained in the email message. The link looks like this: <a href="http://ndpbeta.nla.gov.au/ndp/del/forgotPassword?n=1217895904044053706">http://ndpbeta.nla.gov.au/ndp/del/forgotPassword?n=1217895904044053706</a>

(2) Form presented after user has linked to URL:

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
New password	Password must be between 4 and 20 characters long.
Repeat the new password	Must be the same as the new password.
Reset your password	Button submits the form.

Back to [List of use cases and interface objects](#)

### **Newspaper Titles List**

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Filter by state / territory	Shows the filter applied if any, with option to remove the filter
List of newspaper titles	Title is hyperlinked to <a href="#">Newspaper View</a>
	<p>Only newspaper titles which have searchable articles are included.</p> <p>When a batch is QA3 approved in the QA system, the title will appear in this list if it has not appeared before.</p> <p>Removing a title from this list hopefully will be rare, and removing it from the list will probably remain manual</p>

Back to [List of use cases and interface objects](#)

### **Issues by Date List**

<i>Element on screen</i>	<i>Comments, business rules etc</i>
Standard header	
Date of issue	
Number of results	
List of newspaper titles (which have issues on this date)	Each title is hyperlinked to <a href="#">Page View</a> for the first page of the issue.

Back to [List of use cases and interface objects](#)

### **Edit Immediately or Become a Member**

<i>Element on screen</i>	<i>Comments, business rules etc</i>
General	Window pops up in Article View when user goes to add/edit tags, add/edit comments or correct the text, if the user is not logged in and has not previously done a captcha challenge in this browser session.
Captcha challenge	User must input the letters seen in the captcha. If input is incorrect, an error message is presented and user can try again.
Cancel	Closes the window.
Submit	Submits the completed captcha.
Links to login and signup screens.	Links to <a href="#">Login</a> and <a href="#">Signup</a>

Back to [List of use cases and interface objects](#)

## **Other: Website information ('static' pages)**

These are the pages on the website which contain mostly static data. The text parts of these pages may be edited by editing the corresponding file in the appropriate directory: .

### ***About*** (about.txt)

Dynamic parts of this screen include the graph showing coverage of newspaper titles and dates, and the number of pages and articles available for searching/browsing.

### ***Help*** (searchTips.txt)

### ***Terms of Use*** (termsOfUse.txt)

Back to [List of use cases and interface objects](#)

## **Other: Logs and Statistics**

Kent 28/7/08:

I added Google Analytics stats collection code to the site on Friday afternoon (in the trailer). This means Google Analytics can be used to report usage. You need a google account (gmail address) to be able to access these stats.

Start at <http://www.google.com/analytics/> and after logging in, click the "view reports" link

Kent 4/8/08

.. there's a very tenuous relationship between the Google Analytics stats and stats from our own weblogs.

Google Analytics stats are collected by Google. Google is notified by things happening on the NDP site when and only when a browser going to the NDP site executes some javascript which we include at the bottom of each HTML page on the NDP site. This javascript causes an independent HTTP request to be sent to Google - it counts these and looks at the web page they came from.

Web logs are generated by the web server (jetty) in response to each and every HTTP request regardless of what it is or how it was generated. These logs include everything - every image, every javascript download, every search, whether performed by a real person or

a robot, a web crawler or an availability checker. Robots don't execute javascript, so they don't count in Google Analytic stats.

Back to [List of use cases and interface objects](#)

## **Other: : User Survey**

## **Other: Recent searches**

Most recent searches are accessed with the URL  
<http://ndpbeta.nla.gov.au/ndp/del/search?showRecentSearches=y>

## **Miscellaneous stuff**

Question:

Why does the (heading) OCR seem ok (inaccurate but covering the right bit of text) but the heading zoning seem wrong?

Answer:

The heading zoning provided by APEX is almost certainly correct. The problem with the right white strip is caused by a heuristic weakness (ok, bug) in my code which tries to calculate a "negative" overlay for the page by creating non-overlapping rectangles of "darkness" which cover all the parts of the page not in the article. Because articles typically consist of many zones (parts) on a page (typically, headings and each para are in their own zone), and because the bounding boxes supplied by APEX for these zones typically either don't neatly abut, or even overlap each other, it is an interesting problem to create a neat looking "negative" overlay using as few non overlapping rectangles as possible (and it is important the rectangles don't overlap - doing so would make the page have darker patches where they overlapped).

The current algorithm we are using normally works with a variety of inputs, but is often caught out, as in this case. We have to improve some of our heuristics!

Kent