Requirements

Version 4

August 3, 2007

1 Search

1.1 Query language

1. *Completed.* The user must be able to query by wildcards representing single words, multiple words, a given number of words, multiple characters and single characters

2. *Completed.* The system should provide the user with the ability to use 'or' in queries

1.2 Raw results (before display)

1. *Completed.* For each match found the surrounding text must be returned up to a paragraph or a certain number of words (whichever is smallest)

2. *Completed.* Maximum word limit must be configurable

3. *Completed.* The system should provide the option of "paging" search results at the document level

4. *Completed.* For each match found the matched expression's position (e.g. sentence and paragraph number) in the document must be returned

5. *Completed.* For each match found the document ID of the document containing the matched expression must be returned

6. *Completed.* The returned results must preserve the markup representing non-textual data

7. *Completed.* The user must be able to sort the returned results

   - alphabetically by left/right word up to two levels' depth
   - by position (e.g. sentence and paragraph number) of the matched expression in the document
   - by document ID
   - by the matched expression

2   Filtering

1. *Mostly completed. ETA 10/8.* The user must be able to filter the results by

   - *(Completed)* Discipline
   - *(Completed)* Text type (e.g. essay, report...)
   - *(Completed)* Year of study
   - *(Completed)* Native speaker status
   - *(Completed)* Position in document (i.e. abstract, introduction, middle, conclusion)
   - Paragraph and sentence number

3   System logging

1. *Under assessment.* For each query the system must log

   - The expression to be matched
   - The network address of the user
   - The username of the user if registered
   - The document ID of the matches the user requests further details for
   - Other information required for development purposes

2. *Under assessment.* The system must log the number of concurrent users

3. *Under assessment.* The system must alert of potential automated access

4   User registration

4.1   Anonymous users

1. *Completed.* The system must allow complete access to anonymous users

2. *Under assessment.* After a certain number of searches a prompt inviting the user to register to help the project must be given

3. *Under assessment.* Registration must be completely optional

4.2 Registration process

1. *Under assessment.* When registering the user must be prompted for

   - their valid email address (to act as a username)
   - a password
   - their name (optional)
   - the institution to which they belong (optional)
   - their status (e.g. professor, postgraduate student, undergraduate student etc...) (optional)
   - their feedback (optional)

2. *Under assessment.* The user's email address must be validated before registration is completed

4.3 Registered users

1. *Under assessment.* Registered users must be provided with their search history

2. *Under assessment.* The number of results displayed per page should be configurable by the user

5 User interface

5.1 Accessibility

1. *Completed.* The interface must be web-based

2. *Completed.* The interface must be accessible with a wide range of web browsers on a wide range of platforms and conform to all appropriate standards where practical

3. *Under assessment.* The interface must conform to the "Web Content Accessibility Guidelines" [WACG] W3C standard to priority level 1

4. *Under assessment.* The system must be able support up to 100 concurrent users

5.2 User view

5.2.1 Main page

1. *Completed.* When the user first enters the site the main page must be displayed, displaying introductory information

2. *Completed.* The user must be able to navigate from this page to the Search page

5.2.2 Search page

1. *Completed.* The search page must cleanly and intutively offer options for search functionality

2. *Completed.* The user must be able to specify filters by which to filter the results at this stage

3. *ETA 3/8.* The user must not be able to select contradictory filters (i.e. system must recognise that availability of some filter options is dependent on previously selected filters)

4. *ETA 10/8.* The search page must offer context-sensitive help

5. *Under assessment.* The search must offer the user template queries

6. *Completed.* A successful search must navigate the user to the results page

5.2.3 Results view

1. *Completed.* The results page must display the search results with keywords in context

2. *Completed.* All results must be displayed in tabular form, with matched expressions stacked vertically

3. *Completed.* For each result, the position in the document must be displayed

4. *Completed.* For each result, the document ID must be displayed

5. *Completed.* The user must be able to click on a given result to view the Details view for that result

6. *Completed.* The user must be able to change the sort preference for the displayed results

7. *Completed.* The user must also be able to perform some filtering directly from the results page

8. *Completed.* Any metadata representing non-textual data (e.g. formulae, tables, illustrations...) must be replaced with a simple text subsitute (e.g. [formula], {table} ...)

9. *Completed.* The system should allow access to all documents through document-level paging

10. *Completed.* The system should, by default, return up to 500 results

11. *ETA 10/8.* The system should provide the facility to download search results as a .csv file

12. *Under assessment.* The system should provide statistical information (eg "Your search produced 123 results out of a potential XXX documents [or words]) and this information should be broken down by the filters used

5.2.4   Details view

1. *Completed.* The details view must display an expanded excerpt for the search result in question

2. *Completed.* The document ID of the document must be displayed

3. *Completed.* The location of the matched expression in the document must be displayed

4. *Completed.* The matched expression must be emphasised in some fashion

5. *Completed.* Any metadata representing non-textual data (e.g. formulae, tables, illustrations...) must be replaced with a substitute

5.3   Administrator view

1. *Under assessment.* Administrators must be able to log in to the interface to arrive at administrator view

2. *Under assessment.* Administrators must be able to perform log reporting in administrator view

3. *Under assessment.* Administrators must be able to perform user management in administrator view