# Release Summary

|  |  |
| --- | --- |
| Release | Date |
| 1 | 12/10/2018 |
| 2 | 26/10/2018 |
| 3 | 09/11/2018 |

Release 3 includes:

* Preprocess module
  + Extract web page titles from web pages
  + Optimize the running efficiency and remove redundancy of code.
  + Fix bugs reported in testing
    - Fix the bug that web page title extractor will throw exception when no observer observing it.
* Query module
  + Improve the user interface.
    - User can choose the rank algorithm when ranking the query result.
    - Query result now only shows top 20 pages and shows web page titles.
  + Fix bugs reported in testing
    - Fix the bug that the program will break when the index map does not contain certain URL.

Release 2 includes:

* Preprocess module
  + Remove stop words from keywords set
  + Extract links from web pages
  + Fix bugs reported in testing
    - Fix the bug that generated index map may contains script and style snippets.
* Query module
  + Implement Page rank algorithm to rank the query results
  + Optimize the running efficiency of query module
  + Fix bugs reported in testing
    - Fix the bug that the query process will break when there are concurrent element removal/insertion of query result.

Release 1 includes:

* Preprocess module
  + Generate index map from html web pages.
  + Extract keywords from html keywords tag and main content.
* Query module
  + Implement search function of query module
  + Implement TF-IDF algorithm to rank the query results