**Systems Specification**

**Functional requirements:**

1. All users will access the system using a user name and a password
2. The system shall provide a file upload box, category selection box, and a confirmation button. All users will upload a file in the box, choose a category and click the button, after which users will receive the corresponding text summary.
3. The system shall support the following document formats: PPT, Microsoft Word 2010.
4. The system shall provide a ‘UP’ button and ‘DOWN’ button for each files uploaded. All users will click the ‘UP’ button if the file is considered as worthy reviewing. Otherwise, click ‘DOWN’ button.
5. The system shall provide a search box to enable users to search document by selecting a keyword.
6. All users could download existing files by double-clicking.
7. The system shall provide the user view format of each file.

**Non-functional requirements:**

1. The system shall run as web page which should be compatible on multiple browsers such as IE, Chrome, Firefox etc. The web page shall be fitted to different operating systems including Windows, macOS and Linux-based systems
2. The system shall be programmed in Python (core algorithm).
3. The system shall use Mongo Database to store data.
4. Files uploaded by users shall be stored as tuples in the database.
5. The searching results will be sorted according to the weight for each uploaded document, which will be calculated by an algorithm involved the number of ‘UP’s and ‘DOWN’s.
6. The system shall enforce a password policy: at least 8 characters including at least 1 capital letter, 1 punctuation mark.
7. The system shall not disclose any personal information about users except their log-in information and uploaded files to the operators of the system.
8. The database used in the system shall maintain the ‘ACID’ properties.

**System requirements:**

1. Initial Database Size:

--The initial database shall have the capability to record the information for a minimum of 500 documents. (initially a size of 500 tuples)

1. Rate of Growth:

--The database shall be expandable as files are uploaded to the system.

1. Expected type and frequency of searches:

--The Expected type of system shall be a Python programmed software system including GUI, interface ,core algorithm programs, configuration files [1].

--The frequency of searches depends on how often the users access the system to search target files by entering keywords. The system will hold the capacity of at least ten thousand searches conducted per day.

1. Network and Access requirements

-- The system shall be operated under the condition of network connection.

-- The system shall follows the network protocols which includes Transmission Control Protocol(TCP) and Internet Protocol (IP).

-- The frequent access to database shall be necessary when storing and retrieving users’ data.

1. Performance:

-- The system shall have the capability to support at least 100 users using the web page simultaneously.

-- The system shall take at most 5 seconds to authenticate a user.

1. Security:

-- Encryption methods which follows the Advanced Encryption Standard should be allocated to protect users’ privacy from being stolen.

1. Backup and Recovery:

-- The source code and database of the system shall be backed up every 7 days, and the backup data will be stored in the server disk. The latest 5 copies of backup data will be kept on disk.

1. Legal Issues:

-- The system shall strictly obey the regulations in the UK Data Protection Act 1998.

[1] Sommerville, Ian (2007). ["1.1.1; What is software?"](https://archive.org/details/softwareengineer00isom/page/n29). [Software Engineering](https://archive.org/details/softwareengineer00isom) (8th ed.). [Addison-Wesley](https://en.wikipedia.org/wiki/Addison-Wesley" \o "Addison-Wesley). pp. 5–6. [ISBN](https://en.wikipedia.org/wiki/ISBN_(identifier)" \o "ISBN (identifier)) [978-0-321-31379-9](https://en.wikipedia.org/wiki/Special:BookSources/978-0-321-31379-9" \o "Special:BookSources/978-0-321-31379-9). [OL](https://en.wikipedia.org/wiki/OL_(identifier)" \o "OL (identifier)) [18264252M](https://openlibrary.org/books/OL18264252M).