

Homework 2: Final Data Storage Strategy

❖ Database Structure

➤ Users Database

- U_ID
 - Identifying number of the user used for allocation to files, authentication
- Username
 - User-friendly name that the user can set, saved in database for user authentication
- Password
 - Password of user which is not hashed. If I had more time and if this was a website I would be deploying, I would hash these passwords.
- FirstName
 - The user's first name to display
- LastName
 - The user's last name to display

➤ Files Database

- F_ID
 - Identifying number of that file for easier reference
- Name
 - The title of the file
- Filetype
 - The type of the file (.txt, .pdf, etc.)
- Authors
 - The original creator of the file entered by the user
- Text
 - The text of the TXT, DOCX, or PDF file
- CreationTime
 - The time that the file was originally created, entered by the user

- Source
 - The user ID of the uploader
- Size
 - The size of the file in bytes
- UploadTime
 - The time that the file was uploaded to the database, generated on upload
- Sentiment
 - Object with the text analyzed, and the score and magnitude given to the text by the Google NLP API
- Tags
 - Status
 - ◆ Saves the status of the file, whether that is “Uploaded” or “Analyzed” to display to the user
 - Keywords
 - ◆ Array of keywords given to the text as a result of output from Google NLP API
 - Categories
 - ◆ Array of categories given to the text as a result of output from Google NLP API
- Keys Database
 - Here I store my Flask secret key and my New York Times API Key for access in the app
 - Name
 - Name of the key
 - Key
 - The key itself to use
- Text database was scraped, since text was simply added to files in the files database