

## 1. First Week (17/02/2022 – 23/02/2022)

- Created a rough code that is intended to be a simple search engine
  - o Type of data distance: **Cosine Similarity**
  - o Targetted website is set to kompas.com since the documents are in HTML format
    - Additionally, this is also intended as a starting point before modifying the working code to mine different website(s)
  - o **Known Problems**
    - Bs4 library seems to be missing
    - Test runs are not yet able to be done due to the aforementioned issue

## 2. Second Week (24/02/2022 – 02/03/2022)

- Development\_record created to document details regarding the current progress and upcoming target
- Previously-known issues (in the first week) are not yet fixed, therefore no changes were made to the code

## 3. Third Week (3/02/2022 – 16/03/2022)

- Error involving the bs4 library has been fixed
- Testing the code to kompas.com turned out successful
- Plan to test the search engine to CNBC dataset has been cancelled

## 4. To-do List

- Establish user input for the searched query
- Directing the code to mine PubMed's article abstracts (*final objective*)