SIT322 Cloud Systems: Lab 5 [3 marks]

Prerequisites:

Software

If you are doing this lab on your own computer, you would need the following installed:

- Jdk 1.8
- Eclipse IDE for Java EE Developers https://www.eclipse.org/downloads/packages/

The above are already installed in the university computers in labs T1.01, T1.05& T1.06.

- New to Java?

Teaching Java is out of scope for this unit. However you can find many tutorials online.
https://www.lynda.com/learning-paths/Developer/become-a-java-programmer

- New to Eclipse IDE?

Have a look at these resources BEFORE you start this lab:
https://www.lynda.com/Eclipse-tutorials/Eclipse-Essential-Training/382570-2.html?org=deakin.edu.au
http://www.vogella.com/tutorials/Eclipse/article.html

Task : Upload data to Azure storage

1. In this task, you will upload unstructured text data to an Azure container.
2. Download this folder containing text-based ebooks:
   https://drive.google.com/open?id=1VkKpZclPrIYynqkYTNYKzdD9wkfGUXO8

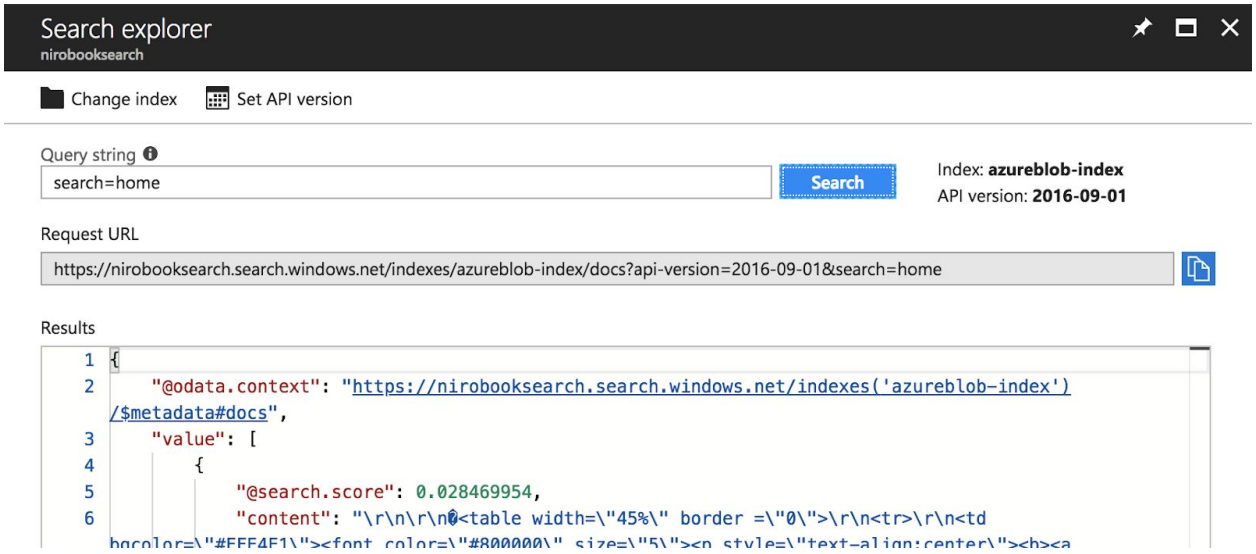
(Note: books are from *Project Gutenberg Australia*)

3. Follow the instructions given in
   https://docs.microsoft.com/en-us/azure/storage/blobs/storage-unstructured-search, but do the following changes:
   a. instead of the **clinical-trials.zip,** upload the given text-based books you downloaded in step 2
   b. When you **Configure the index**, please follow the screenshot given below

Fields

FIELD NAME	TYPE	ATTRIBUTES
content	Edm.String	Searchable, Retrievable
metadata_storage_content_t...	Edm.String	
metadata_storage_size	Edm.Int64	
metadata_storage_last_modi...	Edm.DateTimeOffset	
metadata_storage_name	Edm.String	Filterable, Retrievable
metadata_storage_path	Edm.String	Key, Retrievable
metadata_content_encoding	Edm.String	
metadata_content_type	Edm.String	
metadata_language	Edm.String	

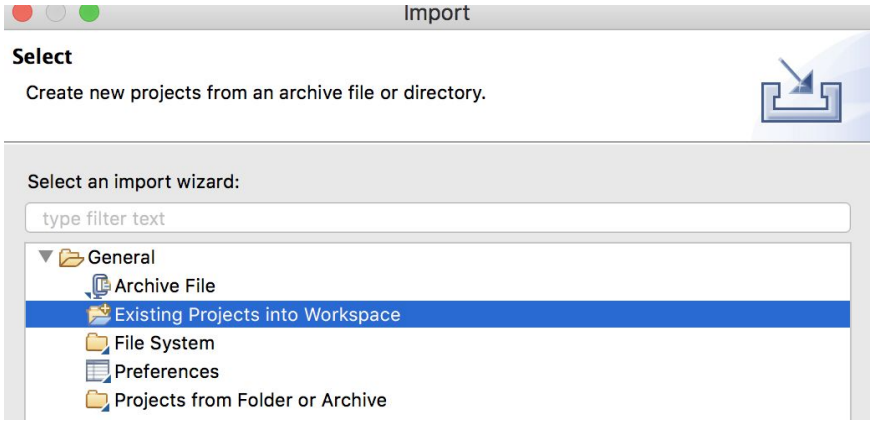
4. Using the graphical ‘Search Explorer’ search for the word ‘home’.



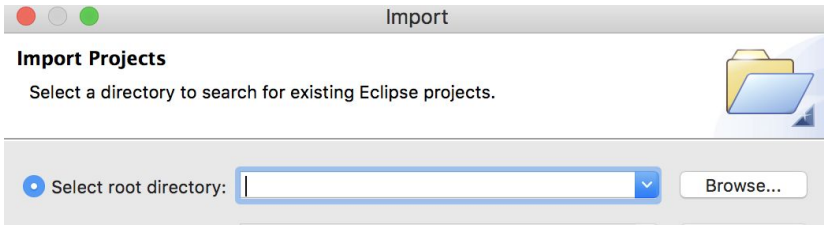
What do you see? Are the results easy to understand?!

5. As you can see, the search results are not very humanly readable. To address this, often you need to write some code to make it more ‘human friendly’. We will do this using Java.
6. Download the following zip file: https://drive.google.com/open?id=1iAKwecLIdsCl4JO15J4q0zxce2uPA0qU

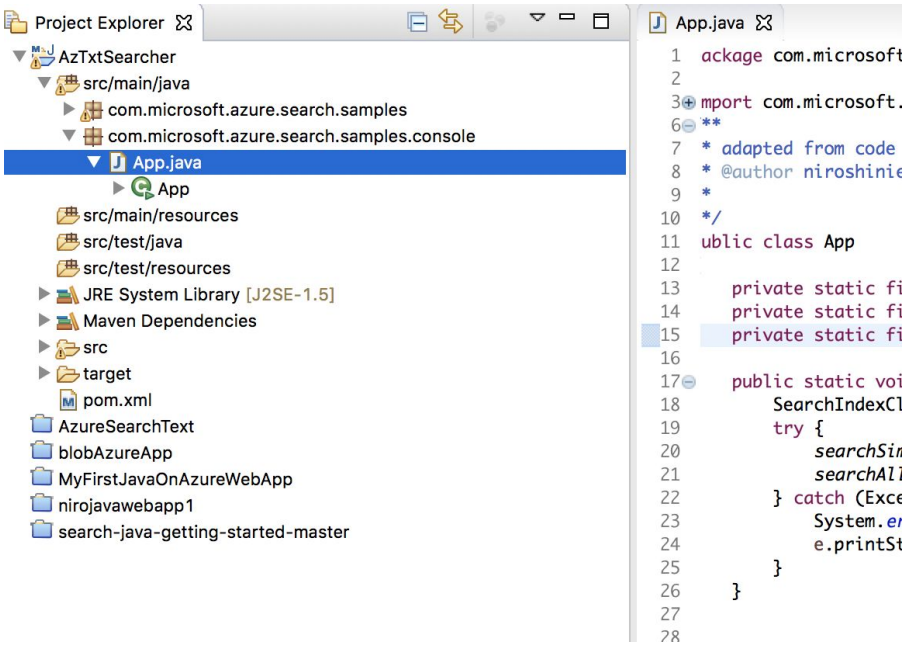
7. Unzip it to your computer. This is an eclipse project. Open your eclipse EE IDE and select File->Import->General->Existing projects in to workspace



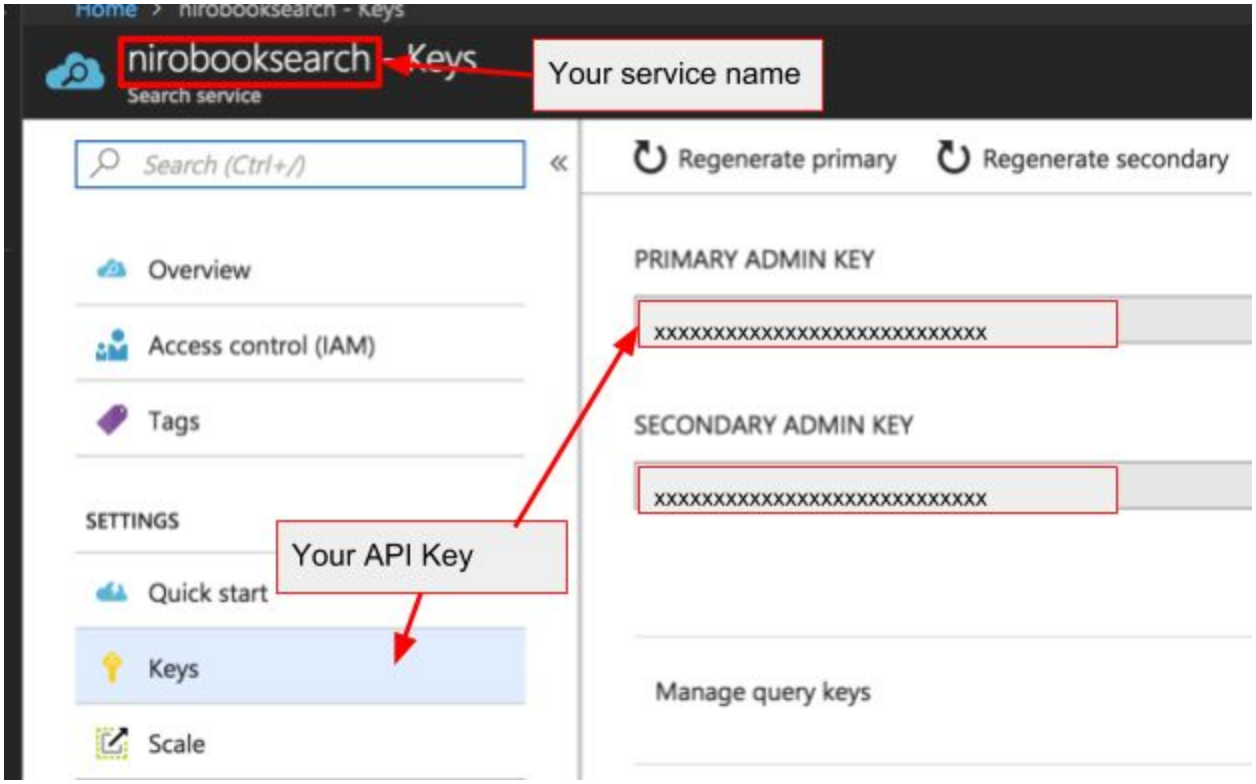
8. Click Browse and select the unzipped project location

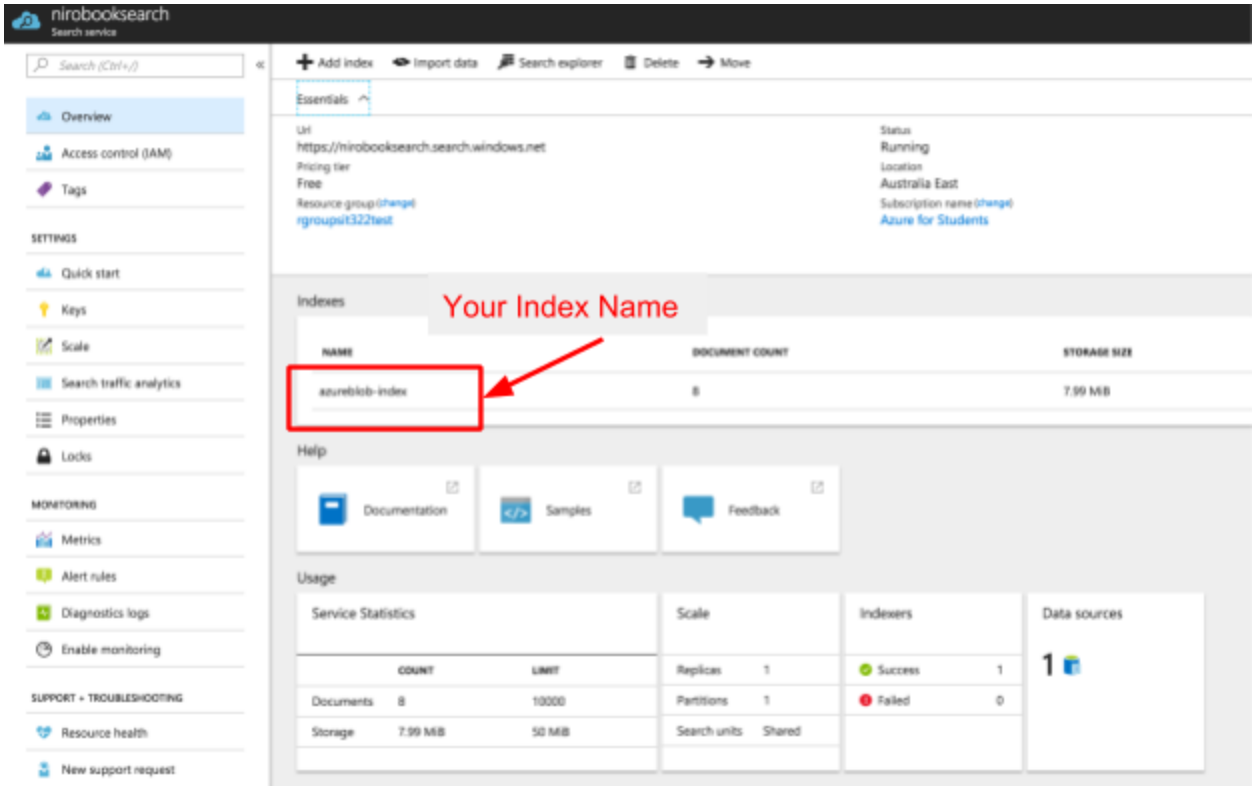


9. Once you have successfully imported the project, you should see the Project Explorer on your left hand side. There, select the App.java file located under \src\main\java\com\microsoft\azure\search\samples\console and double click to open it.



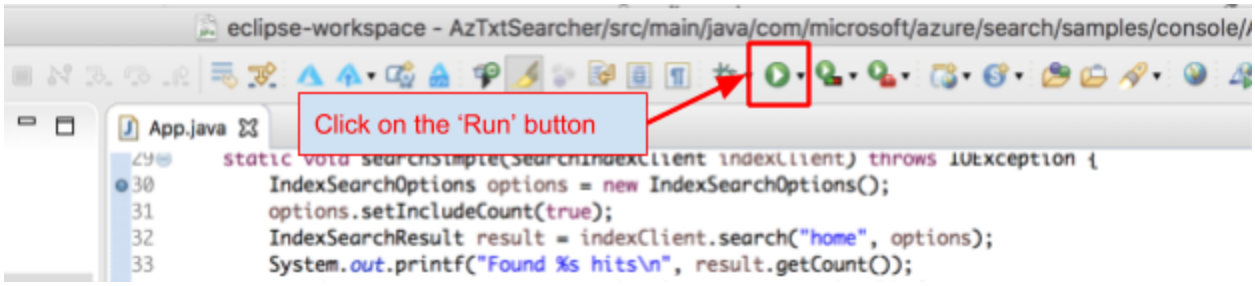
10. Inside the App.java file update the three variables SERVICE_NAME, INDEX_NAME & API_KEY with values from your created Azure search.


nirobooksearch - Keys
Search service
Your service name
Search (Ctrl+/)
Regenerate primary
Regenerate secondary
Overview
Access control (IAM)
Tags
PRIMARY ADMIN KEY
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
SECONDARY ADMIN KEY
xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
SETTINGS
Your API Key
Quick start
Keys
Manage query keys
Scale


nirobooksearch
Search service
Add index
Import data
Search explorer
Delete
Move
Essentials
https://nirobooksearch.search.windows.net
Free
Running
Australia East
Azure for Students
Indexes
Your Index Name
NAME
DOCUMENT COUNT
STORAGE SIZE
azureblob-index
7.99 MB
Help
Documentation
Samples
Feedback
Usage
Service Statistics
Scale
Indexers
Data sources
Overview
Access control (IAM)
Tags
Quick start
Keys
Scale
Search traffic analytics
Properties
Locks
Metrics
Alert rules
Diagnostics logs
Enable monitoring
Resource health
New support request

11. Have a look at the code in App.java. In the main method, we are calling two methods: **`searchSimple`** & **`searchAllFeatures`**

12. Run the code



Task Submission:

Answer the following questions in a document -

A)

1. In the given code, what is the word being searched? (0.25 marks)
2. Run the code and get a screenshot of the eclipse console. Include the screenshot in your answer sheet. (0.25 marks)
3. Explain the console output. (0.5 marks)

B)

1. Change the code to search for your student id (change both methods) and include he screenshot the eclipse console in your answer sheet. (0.25 marks)
2. Change the code to search for the word "why". You must do both of the following:
    a. Search for 'why' in all of the books
    b. Apply a filter to search for 'why' in the following books only - 'FarFromTheMaddingCrowd.txt', 'LittleMen.txt'

    - For (a) & (b) Include the screenshot of the eclipse console in your answer sheet. (1 mark)
    - Also submit your modified App.java file (0.5 marks)
3. Imagine you need to add another book to your storage and run the search again. Explain what you should do to get an accurate search result after you add the new file. (0.25 marks)

.

Upload the document as a PDF and your modified App.java file to the relevant dropbox on Cloud Deakin by specified deadline. Please refer to Cloud Deakin dropbox for due dates.

References:

https://docs.microsoft.com/en-us/azure/storage/blobs/storage-unstructured-search
https://docs.microsoft.com/en-us/azure/search/search-video-demo-tutorial-list
https://docs.microsoft.com/en-us/azure/search/search-get-started-portal