2021

Twitter Enalyst

CAB432

Assignment 2

Dac Duy Anh Nguyen n10603280

Do Le Hoang Minh n00000000

Contents

[Introduction 2](#_Toc52357527)

[Purpose & description 2](#_Toc52357528)

[Services used 2](#_Toc52357529)

[Twitter Standard Search API (v.1.1) 2](#_Toc52357530)

[API 2 2](#_Toc52357531)

[API n 2](#_Toc52357532)

[Use cases 2](#_Toc52357533)

[US 1 2](#_Toc52357534)

[US 2 2](#_Toc52357535)

[US n 3](#_Toc52357536)

[Technical breakdown 4](#_Toc52357537)

[Architecture 4](#_Toc52357538)

[Context diagram 4](#_Toc52357539)

[Sequence Diagram 4](#_Toc52357540)

[Process flow Diagram 5](#_Toc52357541)

[Network diagrams (Cloud specific) 5](#_Toc52357542)

[Client / server demarcation of responsibilities 5](#_Toc52357543)

[Response filtering / data object correlation 5](#_Toc52357544)

[Test plan 7](#_Toc52357545)

[Difficulties / Exclusions / unresolved & persistent errors 8](#_Toc52357546)

[Extensions (Optional) 8](#_Toc52357547)

[User guide 8](#_Toc52357548)

[References 8](#_Toc52357549)

[Appendices 8](#_Toc52357550)

This template is similar to the one provided for assignment 1. It is not compulsory to use it, but it will save a lot of effort if you do. You should assume that black text in italics is there as guidance and you should read it, follow the instructions and then delete it when you have entered your own text. Some examples are not italicized, but should obviously be replaced by your own material.

The report should be around 10-15 pages including screenshots, but this is a guide only – we will not be enforcing a page limit or marking you down for submitting something with 16 pages instead. But be sensible, we really don’t want something that is 25 pages or more. Please note there are examples of previous students’ work in this template & examples from Google searches. They are here to give you ideas on what you can do. **We do not guarantee that they are appropriate for your project.**  You must do your own research and produce your own diagrams.

## Introduction

### Purpose & description

Twitter Enalyst is an amazing application for anyone interested in discovering and analyzing sentiment factor of certain topics. And Twitter Enalyst allows users to search for a keyword and to receive sentiment score of the most recent posts containing that keyword. The app also suggests trending keywords on Google trends at that moment which users can test their sentiment score on our app. On the server side, this app utilizes Sentiment module to perform sentiment analysis on input Twitter posts that are queried from Twitter API.

Ss (screenshot)

### Services used

#### Search Tweets: Standard API v.1.1

Returns a collection of relevant Tweets matching a specified query – can also be filtered based on time, location, language, geography, etc.

Endpoint: <https://api.twitter.com/1.1/search/tweets.json>

Docs: <https://developer.twitter.com/en/docs/twitter-api/v1/tweets/search/api-reference/get-search-tweets>

#### Sentiment API

A node.js module that uses the AFINN-165 wordlist and Emoji Sentiment Ranking to perform sentiment analysis on arbitrary blocks of input text.

Endpoint: using from library

Docs: <https://www.npmjs.com/package/sentiment>

#### Google Trend API

A node.js module that providing trending keywords from

Endpoint: using from library

Docs: <https://www.npmjs.com/package/google-trends-api>

### Use cases

#### User story 1

|  |  |
| --- | --- |
| As a | User |
| I want | To see what keywords are trending right now in Australia on google search engine |
| So that | I may use them as input for sentiment analysis |

For this user story, we implement the Google Trend library to get trending keywords in Australia each time a new user access our app. The server is responsible for requesting data. After the keywords is retrieved, they will be sent to client side by how….

Ss of client side, show code that does this

#### User story 2

|  |  |
| --- | --- |
| As a | User |
| I want | To search for a keyword |
| So that | The app can return the sentiment score of posts containing that keyword within the last 7 days |

On the client side, users can type in their desired keyword to analyze sentiment score. Once they click ‘Search’, that keyword will be sent to the server via route “/abc” to start sending requests to Twitter API and analyze responses’ text. The results will be sent back to client side when they are ready.

Ss of the search bar

#### User story 3

|  |  |
| --- | --- |
| As a | User |
| I want | The app to display a chart of sentiment score summary (total positive and negative scores) |
| So that | The result is more intuitive to users |

As said in the previous user story, the analysis results received by client side is handled by d3.js to create chart with visualized result summary, namely Total positive/negative score, Number of positive/negative/neutral posts.

Ss sentiment score chart, show code

#### User story 4

|  |  |
| --- | --- |
| As a | User |
| I want | The post analyzed to be displayed on the app |
| So that | I can know what those posts says |

Minh

Ss, show code that does this

## Technical breakdown

In this report – which covers the group components of the assignment - there should be some coverage of the architecture and the basic operation of the system. Some deeper analysis is now left to the individual report – please read that template and guide for details.

### Architecture

Explain how your system operates, making it clear how data flows around the system through requests and responses, and the appearance of scaling and persistence within the architecture. In this report it is not necessary to discuss in detail the effect of these choices. Here we just want you to document the architecture and to tell us how it works.

Your principal helper here will be one or more architecture diagrams – which we will consider in some more detail below. However, you may show us screen grabs of code if that makes your points clearer. Tell us anything you think we need to know about how you have structured the application and made it work, but there also a section below to describe problems. If you have used particular libraries, then you should give us a brief overview of their use in this application.

A number of example architecture diagrams are provided below. Many students use diagram generators such as the tools at <https://cloudcraft.co/>. For assignment 2, this is the most important diagram used to document your approach. The ‘network diagrams’ below show some more complicated alternatives. Only the architecture diagram is compulsory. Please consider the others if they help you, and ignore them if they do not. Obviously you should delete all of the examples and include only diagrams which you have created to explain your application.

#### Context diagram

|  |  |
| --- | --- |
|  |  |

#### Sequence Diagram

|  |  |
| --- | --- |
|  | C:\Users\denbi\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\5B3F0D8D.tmp |

#### Process flow Diagram

|  |  |
| --- | --- |
|  |  |

#### Network diagrams (Cloud specific)

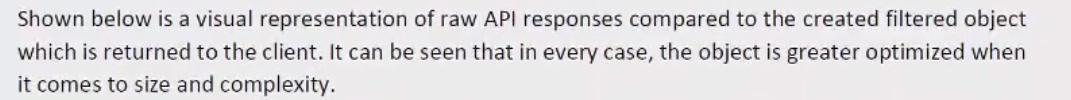
|  |  |
| --- | --- |
|  |  |

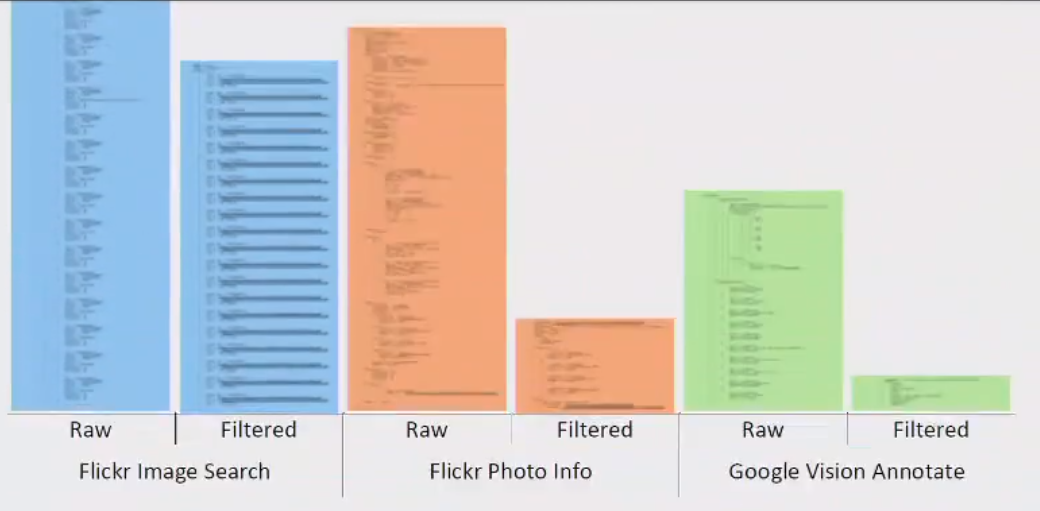
#### Client / server demarcation of responsibilities

Explain to us what is doing what & where. Refer to the architecture diagram and any others that you find appropriate. This is particularly effective if you support your comments with well-chosen code fragments. These should be short and focused and you should give us any context that we need to work with them.

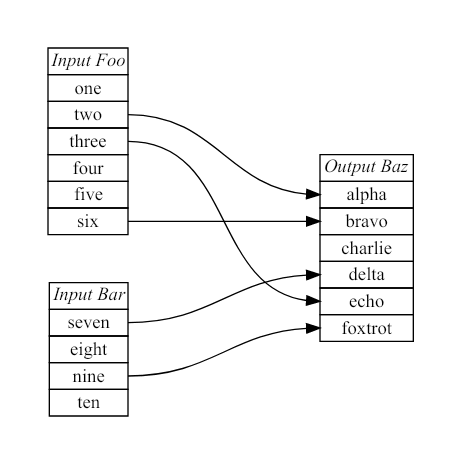
#### Response filtering / data object correlation

Show us how you manipulated the data. The same comments apply about referring to the diagrams and supporting your work with code fragments as appropriate. We provide an example below of how to show this diagramatically. Please note that this example is quite specific to the system being explained. Yours might have an entirely different look, but do a similar job. This section will vary markedly according to the application and may not make much sense for some custom applications such as rendering.





You could instead use a data relationship diagram:



### Scaling and Performance

This is a crucial aspect of the report, and you should use this section to document the approach taken to scaling – the nature of the application load, how it was varied and how the scaling infrastructure responded. You should refer to the architectural diagram above or reproduce the relevant aspects here. You should include screenshots of CPU, network or queuing metrics as observed on the cloud services dashboard, together with screenshots of your settings and the scaling pool instance creation and destruction. We expect that your work here will demonstrate successful scale out and scale in as required in the assignment specification. The screen shots that you use here will also very likely be re-purposed as part of the slide deck for the demo.

An example scaling image is shown below, and we would normally expect to see this sort of image and some evidence of your group settings. Note the instance count on the left hand axis. As noted, many alternatives are possible.



### Test plan

Manual testing is fine and our expectations are in line with the example grid below. You can show the results through a screen shot and point us to these from the table.

Your tests should include

* Positive outcome cases
* Negative outcome cases (error scenarios)
* Edge cases
* Non-functional cases

|  |  |  |  |
| --- | --- | --- | --- |
| Task | Expected outcome | Result | Screenshot/s |
| Data stored in Redis | New keyword result is stored in Redis |  |  |
| Data stored in DynamoDB | New keyword result is stored in DynamoDB |  |  |
| Retrieve Data from DynamoDB | Result of repeated keyword within 24 hours is retrieved from DynamoDB |  |  |
| Retrieve Data from Redis | Result of repeated keyword within 24 hours is retrieved from DynamoDB |  |  |
| Retrieve Data from Twitter API | New keyword result is queried directly from Twitter API |  |  |
| Trending keyword is got from Google Trending API | New trending keywords are displayed on client side |  |  |
|  |  |  |  |

As they are common in industry you could define your Acceptance Criteria as GWT statements. This is not compulsory, but see: <https://www.agilealliance.org/glossary/gwt/>. And here is an example:



Difficulties / Exclusions / unresolved & persistent errors /

Could not scale with using Stream Twitter API and Socket - Minh

Variable declaration type - Rodo

In this section, you should explain anything that caused you problems and how you overcame those problems. Tell us if there was any issue that prevented you completing the assignment to specification. Tell us about any assumptions or compromises that you have made. Those who worked with an API like Spotify, which presented particular concerns, should discuss the compromises here, and this is also where you can tell us about problems with API keys and responses.

More generally, you might consider:

* Your major roadblocks and how you resolved them.
* Any functionality you didn’t or couldn’t finish
* Are there any differences between your brief and what you delivered? If so, explain why.
* Are there any outstanding bugs?

## Extensions (Optional)

In this section, you can tell us if you wish to how you might extend your app and make it better. This is an opportunity to tell us about good ideas that you had that you didn’t have time to tell us about.

In order to increase the precision of the sentiment scores, we could increase the number of posts and the time range queried from API and. At the moment, due to the limitation of free tier access to the Twitter API, we can only query 100 posts that lies with 7 days period. With a small number of posts in a short time range being analyzed, the sentimental scores do not reflect exactly all the twitter posts in the past 7 days from the moment of query.

## User guide

## References

## Appendices