Logo

Description automatically generated

Report 2: Prototype IR System

By

Mr. WARIS DAMKHAM 6388014

Mr. THANAWATH HUAYHONGTHONG 6388016

Miss PATTANAN KORKIATTRAKOOL 6388022

**Submitted to  
Assoc. Prof. Dr. Suppawong Tuarob**

A Report Submitted in Partial Fulfillment of

the Requirements for

ITCS 414 – Information Retrieval and Storage

Faculty of Information and Communication Technology

Mahidol University

2021

**Table of Contents**

[**Introduction** 1](#_Toc120854416)

[**Problem(s) that you are trying to solve** 2](#_Toc120854417)

[**Existing relevant systems** 3](#_Toc120854418)

[**Sea Me Like I See You Search System Implementation** 4](#_Toc120854419)

[**Data collection** 4](#_Toc120854420)

[**Example documents** 4](#_Toc120854421)

[**Data statistics** 5](#_Toc120854422)

[**Tools and software** 5](#_Toc120854423)

[**System diagram** 6](#_Toc120854424)

[**Snapshots of the system** 7](#_Toc120854425)

[**Search Sessions Examples** 8](#_Toc120854426)

[**One word query** 8](#_Toc120854427)

[**Multiple word query** 8](#_Toc120854428)

[**Partial match** 9](#_Toc120854429)

[**Ranking** 9](#_Toc120854430)

[**Discussion** 10](#_Toc120854431)

[**Limitations of your system** 10](#_Toc120854432)

[**Technical difficulties, challenges, and lessons learned** 10](#_Toc120854433)

[**Opportunities for future improvements** 10](#_Toc120854434)

[**Conclusion** 11](#_Toc120854435)

# **Introduction**

# **Problem(s) that you are trying to solve**

# **Existing relevant systems**

# **Sea Me Like I See You Search System Implementation**

## **Data collection**

ไต๋เขียน

Example of data document

## **Example documents**

Keep Common name, Scientific name, Order and Family, Image, and Description

**Data statistics** The JSON data is loaded as the sea index into the Elasticsearch search engine. The JSON data file contains 137 records, which are analyzed and translated into 137 documents for the marine index.

Graphical user interface, text, application

Description automatically generated The data in each document is composed of mainly five fields: the marine animal's common name ("Common name"), the image of the marine animal in jpg and png format ("img-src"), the description of the marine animal ("Description"), the marine animal's scientific name ("Scientific name"), and the order and family of the marine animal ("Order and Family").

## **Tools and software**

* **Elasticsearch**

We have utilized Elasticsearch as our search engine. The following are the primary settings that we have implemented.

1. For ranking, we have used the ascending (smallest value first) sorting module, which is the Standard module. Therefore, no configuration was needed.

"sort": [

            { "Common name": "asc" }

        ]

1. For the data analyzer, we have used a custom Elasticsearch analyzer, which is the implementation analyzer. We have configured this in Elasticsearch’s development tool. The configuration is as follows.

A picture containing timeline

Description automatically generated

* **Flask**

Flask was chosen as our web service to facilitate the use of Elasticsearch as our search engine.

## Diagram Description automatically generated**System diagram**

## **Snapshots of the system**

Graphical user interface, website

Description automatically generated Our search system’s websites are made of two pages: the inde2.html page and the search.html page. The page screenshots are displayed below.

Figure 5 shows the search page of Sea Me Like I See You search system

Text

Description automatically generated(inde2.html page)

Figure 6 shows the search results page of Sea Me Like I See You search system

(search.html page)

## **Search Sessions Examples**

### **One word query**

Example query is “Shark”

Text

Description automatically generatedThe search result is the document that contains the term “Sharks”.

### **Multiple word query**

Text

Description automatically generatedExample query is “large whales” The top two search results are the documents that contain the terms “large “and “whales”.

### **Partial match**

Example query is “large Great shark”

Text

Description automatically generatedThe top two search results are the documents that contain the terms “large” and “shark”, but do not contain the term “great”.

### **Ranking**

Example query is “Sharks”

**Text

Description automatically generated**The first search result contains the word "Sharks" and is therefore very relevant to the search query. This search result is sorted by ascending order in addition to the common name.

# **Discussion**

## **Limitations of your system**

## **Technical difficulties, challenges, and lessons learned**

## **Opportunities for future improvements**

# **Conclusion**