

Faculty of ICT, Mahidol University

Project 2: Prototype IR System

by

6488004 Kittipich Aiumbhornsin 6488089 Pattaravit Suksri 6488092Thanaphat Sumrith

Submitted to

Subject ID: ITCS414

Subject Name: Information Storage and Retrieval

Instructor: Assoc. Prof. Dr. Suppawong Tuarob

A report submitted as the partial fulfillment of the requirements for the project $Semester\ 1/2023$

Introduction

This report describes the creation of a Full Text Food Search Engine as part of a project for Information Storage Retrieval. The Elasticsearch and Flask-powered search engine is intended to give extensive food-related search capabilities.

Problem

This project's major goal is to efficiently retrieve specific food-related information from a large and diverse dataset. There is a major demand for a targeted search tool that can give consumers with exact and contextually relevant results in a user-friendly manner in the digital age, where an overwhelming amount of culinary content is available online.

Existing Relevant Systems

General-purpose search engines and other Elasticsearch-based applications are examples of comparable systems. Our Full Text Food Search Engine, on the other hand, distinguishes itself by focusing solely on the culinary domain, with the goal of providing more refined and relevant search results for food-related inquiries.

Implementation

Our Full Text Food Search Engine was implemented using Elasticsearch and Flask. Data was collected from the BBC Good Food website, and we organized into an Excel table before being exported as (ND)JSON format for indexing in Elasticsearch. The system utilizes a user-friendly interface to accept search queries and displays relevant results based on keyword relevance as we use the default TF-IDF setting algorithms. The engine can handle one-word queries, multiple-word queries, and partial matches, providing users with comprehensive food-related information.

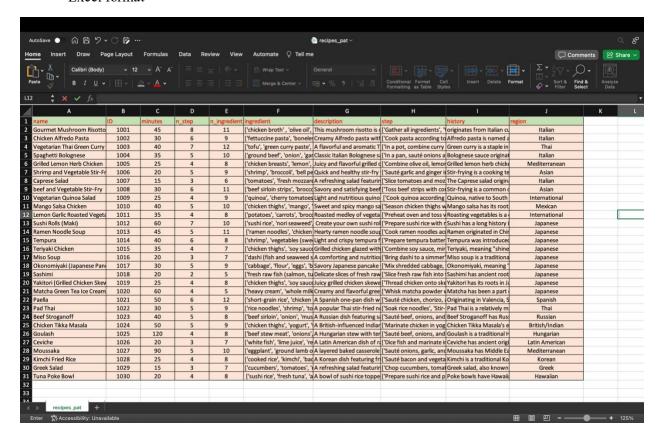
Tools and Software

- Elasticsearch: To provide powerful full-text search capabilities.
- Flask: As a Python based lightweight web application framework to manage the user interface and server-side logic.

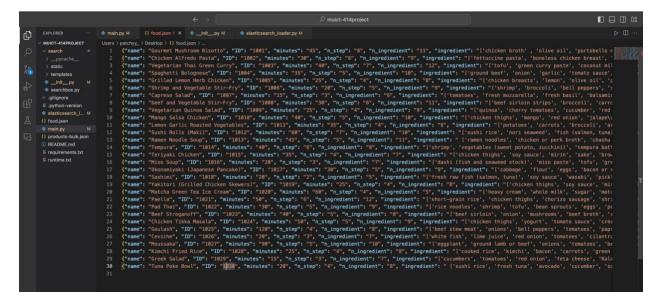
Source of Data

We've gathered data from "https://www.bbcgoodfood.com/recipes/collection/easy-recipes" organized it into an Excel table, and then exported it as JSON format to enhance search functionality in Elasticsearch.

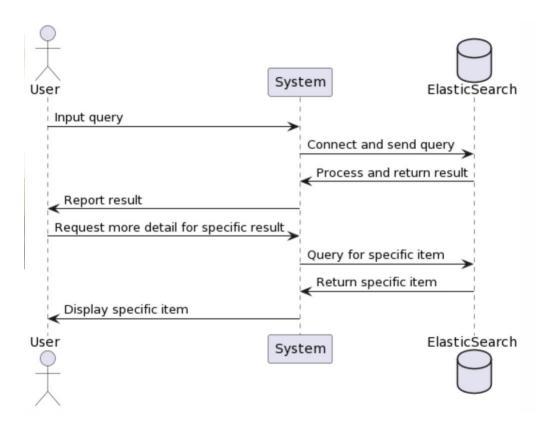
Excel format



• (NDJSON) .json format



System Diagram



Explanation:

- System receives query from user. System connected and send the query to the elastic search.
- Elastic searches use the user's query as an input, then elastic search will process and return the search result
- System, then will report the result of the elastic server to display out the corresponding result(s).
- If user requested for the detailed result for the specific result, system will query for the specific item and then return to the user.

Snapshots of the System

• Search Page



• One word query (rice)

Search Results for "rice"

Page item order	Title	Descriptions	Region	Details
1	Paella	A Spanish one-pan dish with a rich combination of rice, seafood, and meats.	Spanish	link
2	Kimchi Fried Rice	A Korean dish featuring fried rice with kimchi, vegetables, and a savory sauce, often topped with a fried egg.	Korean	link
3	Matcha Green Tea Ice Cream	Creamy and flavorful green tea ice cream made with matcha powder.	Japanese	link
4	Tuna Poke Bowl	A bowl of sushi rice topped with fresh tuna, avocado, cucumber, and a flavorful soy-sesame dressing.	Hawaiian	link
5	Sushi Rolls (Maki)	Create your own sushi rolls with a variety of fresh ingredients.	Japanese	link



For more information you can click "Link" and it will show you more information

KIMCHI FRIED RICE

Data ID: 1028 Name: Kimchi Fried Rice

Description

A Korean dish featuring fried rice with kimchi, vegetables, and a savory sauce, often topped with a fried egg.

History

Kimchi is a traditional Korean side dish of fermented vegetables. Kimchi fried rice is a modern creation, combining Korean and global flavors.

Region

Korean

Estimated cooking time (minutes)

25

Number of cooking steps

1

Number of ingredients

8

Ingredients

['cooked rice', 'kimchi', 'bacon', 'carrots', 'green onions', 'soy sauce', 'sesame oil', 'fried eggs']

Steps

['Saut] bacon and vegetables in sesame oil', 'Add cooked rice and kimchi', 'Stir in soy sauce', 'Top with green onions and a fried egg']

• Multi-word query (Japanese food)

Search Results for "Japanese food"

Page item order	Title	Descriptions	Region	Details
1	Yakitori (Grilled Chicken Skewers)	Juicy grilled chicken skewers with a sweet and savory glaze.	Japanese	link
2	Yakitori (Grilled Chicken Skewers)	Juicy grilled chicken skewers with a sweet and savory glaze.	Japanese	link
3	Okonomiyaki (Japanese Pancake)	Savory Japanese pancake with a variety of toppings.	Japanese	link
4	Okonomiyaki (Japanese Pancake)	Savory Japanese pancake with a variety of toppings.	Japanese	link
5	Chicken Alfredo Pasta	Creamy Alfredo pasta with tender chicken breast, a perfect comfort dish.	Italian	link



For more information you can click "Link" and it will show you more information

YAKITORI (GRILLED CHICKEN SKEWERS)

Data ID: 1019

Name: Yakitori (Grilled Chicken Skewers)

Description

Juicy grilled chicken skewers with a sweet and savory glaze.

History

Yakitori has its roots in Japanese street food culture and has evolved into a popular izakaya (pub) dish.

Region

Japanese

Estimated cooking time (minutes)

25

Number of cooking steps

4

Number of ingredients

8

Ingredients

['chicken thighs', 'soy sauce', 'mirin', 'sake', 'sugar', 'green onions', 'sesame seeds', 'lemon wedges']

Steps

[Thread chicken onto skewers', 'Mix soy sauce, mirin, sake, and sugar for the glaze', 'Grill skewers, basting with the glaze', 'Sprinkle with chopped green onions and sesame seeds, serve with lemon wedges']

• Partial match ("sala" for "salad")

Search Results for "sala"

Page item order	Title	Descriptions	Region	Details
1	Mango Salsa Chicken	Sweet and spicy mango salsa served over grilled chicken thighs.	Mexican	link
2	Mango Salsa Chicken	Sweet and spicy mango salsa served over grilled chicken thighs.	Mexican	link
3	Caprese Salad	A refreshing salad featuring tomatoes, fresh mozzarella, and basil drizzled with balsamic glaze.	Italian	link
4	Greek Salad	A refreshing salad featuring crisp vegetables, feta cheese, olives, and a drizzle of olive oil.	Greek	link
5	Caprese Salad	A refreshing salad featuring tomatoes, fresh mozzarella, and basil drizzled with balsamic glaze.	Italian	link



New Search

For more information you can click "Link" and it will show you more information

GREEK SALAD

Data ID: 1029 Name: Greek Salad

Description

A refreshing salad featuring crisp vegetables, feta cheese, olives, and a drizzle of olive oil.

History

Greek salad, also known as Horiatiki, has been a staple in Greek cuisine for centuries, showcasing the region's emphasis on fresh and simple ingredients.

Region

Greek

Estimated cooking time (minutes)

15

Number of cooking steps

3

Number of ingredients

7

Ingredients

['cucumbers', 'tomatoes', 'red onion', 'feta cheese', 'Kalamata olives', 'olive oil', 'oregano']

Steps

['Chop cucumbers, tomatoes, and red onion', 'Combine with olives and feta cheese', 'Drizzle with olive oil, sprinkle with oregano']

Example Search Sessions

- 1. One Word Query: Demonstrate a search for a term like "rice" and its results.
- 2. Multiple Word Query: Show a search for a phrase, such as "Japanese food"
- 3. Partial Match: Illustrate the system's response to a partial query like "sala" for "salad".
- 4. Ranking: Keyword relevance, Elasticsearch ranks based on relevance to the search query. This is calculated by algorithms like TF-IDF (Term Frequency-Inverse Document Frequency) which consider frequency of query terms.

Limitations

While the system can handle the query and retuned the search result, it may be limited in its ability to handle inquiries in languages other than English or non-text information. Furthermore, the specificity of user inquiries may affect the accuracy and relevance of search results.

Technical Challenges and Lessons Learned

The project offered difficulties in understanding the complexities of Elasticsearch, the significance of dataset, and the application logic and interface. These challenges were all significant learning experiences.

1. Integration with Elasticsearch:

- Challenge: Elasticsearch installation and configuration for efficient data indexing and retrieval. Handling connection problems.
- Lesson Learned: Understanding Elasticsearch setup and security factors is critical for robust and effective search capabilities.

2. Error Handling and Debugging:

- Challenge: Putting in place robust error handling methods, particularly for dealing with search requests and Elasticsearch interactions. Troubleshooting Flask routing and Elasticsearch queries. To be more specific, to showing the detailed page for each query result facing an unexpected view and unable to link and show the specific detail of the search result to show.
- Lesson Learned: Learned the importance of thorough error handling and debugging procedures in ensuring system reliability, applying techniques and knowledge for a real world-like scenario functionality.

Future opportunity

A compelling prospect for future enhancement involves the incorporation of advanced machine learning techniques to continually refine the relevance algorithms employed by the Full Text Food Search Engine. Through machine learning, the system has the potential to adapt and improve based on user interactions, preferences, and behaviors, offering increasingly personalized and contextually accurate search results as users engage with the platform. This dynamic adaptation not only ensures a more tailored and user-centric experience but also positions the search engine to stay agile in response to shifting culinary trends and evolving user expectations. In doing so, it solidifies its role as an indispensable tool for individuals seeking precise and personalized food-related information in the ever-evolving landscape of digital content.

Conclusion

This report summarizes the key components of your Full Text Food Search Engine project. More specific information, especially in the 'Data Collection and Statistics' and 'Technical Challenges and Lessons Learned' sections, as well as the real system diagram and interface snapshots.