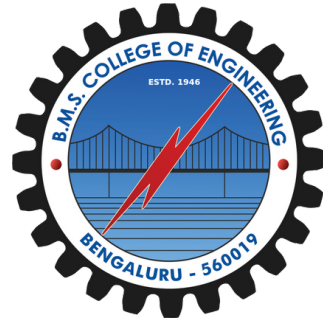




**Hewlett Packard**  
Enterprise



# Open Source Finder

## Student Team:

Akshat Pandey (IBM20IS015)  
Arpan Bhusal (IBM20IS198)  
Gagandeep N K (IBM20IS038)  
Kshama Bhatt (IBM20IS202)  
Rahul T G (IBM20IS112)  
Rohit DB (IBM20IS122)

## Faculty Mentor:

Dr. Sindhu K  
Prof. Nalina V

## HPE Mentor:

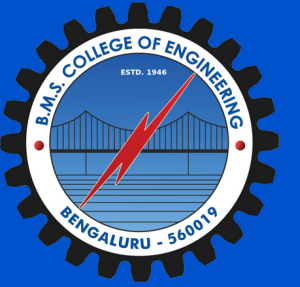
Mr. Arun Ramachandra  
Mr. Murali Krishna

# Agenda

- Introduction
- Abstract
- Initial Architecture
- APIs – Libraries.io
- Implementation of Backend API  
using Flask
- Elastic Search – Fuzzy Logic
- Demo
- Learnings
- Next steps – (Novel Idea)



**Hewlett Packard**  
Enterprise

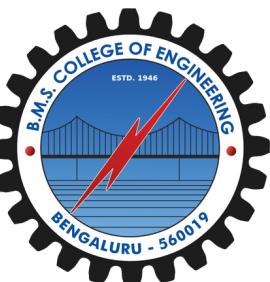


# Introduction

- Open source software is chosen because it can be easily modified or customized to meet a business problem, and it is highly stable.
- Using OSS is considerably cheaper than building software from scratch.
- Choosing an OSS from a considerably large pool of OSS available in the market, we need to look into various factors that takes up lot of time and effort.
- Our project is a solution to this problem. The system designed will present the appropriate OSS required based on one's needs



**Hewlett Packard**  
Enterprise

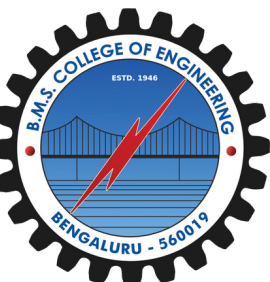


# Abstract

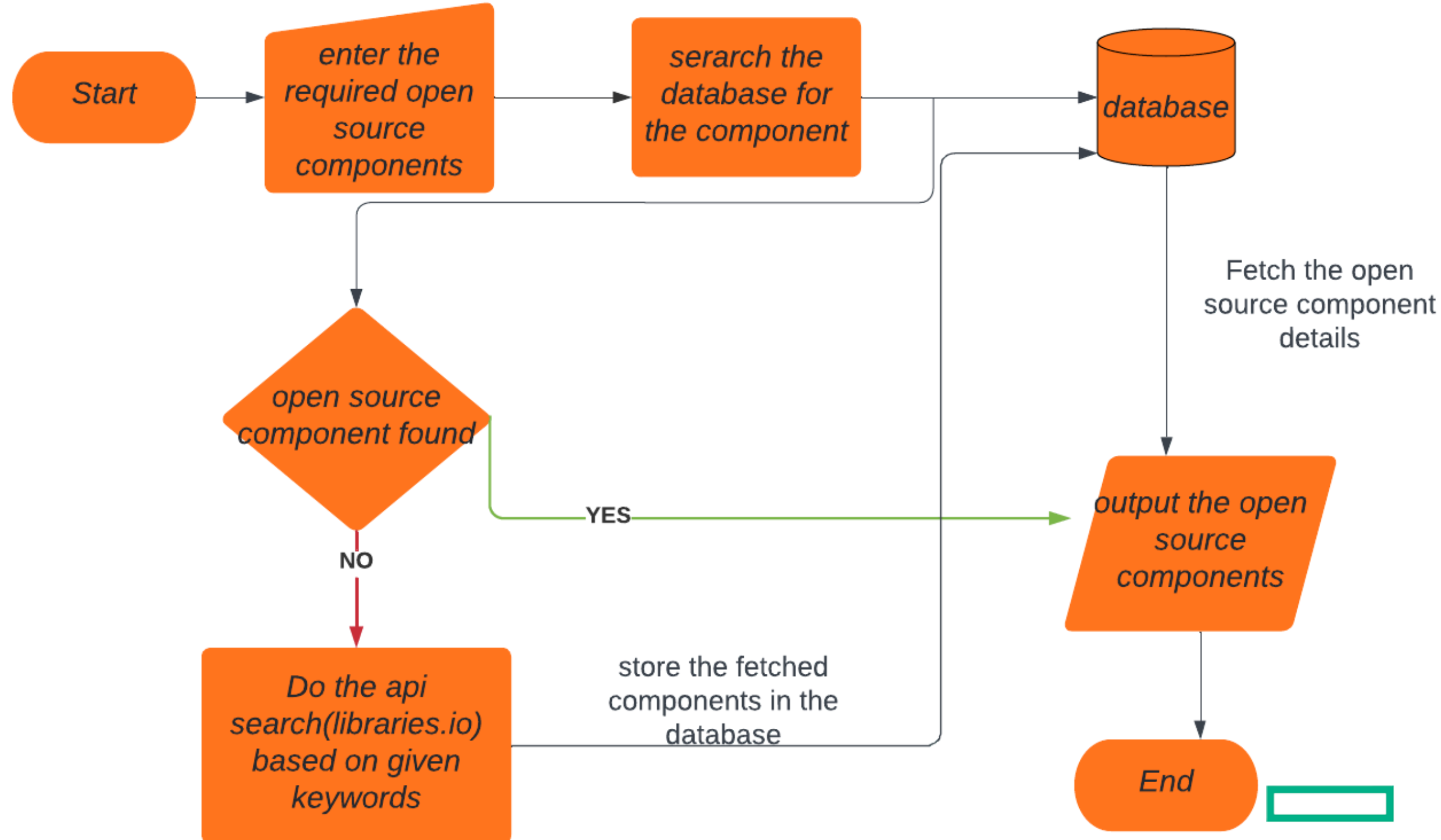
- Open Source Finder is an open-source application that allows users to search for open-source components for their projects.
- It provides a user-friendly interface to discover and explore various open-source repositories based on specific criteria such as license and platform.
- With Open Source Finder, developers can easily find and integrate open-source components into their projects, saving time and effort.



**Hewlett Packard**  
Enterprise



# Initial architecture



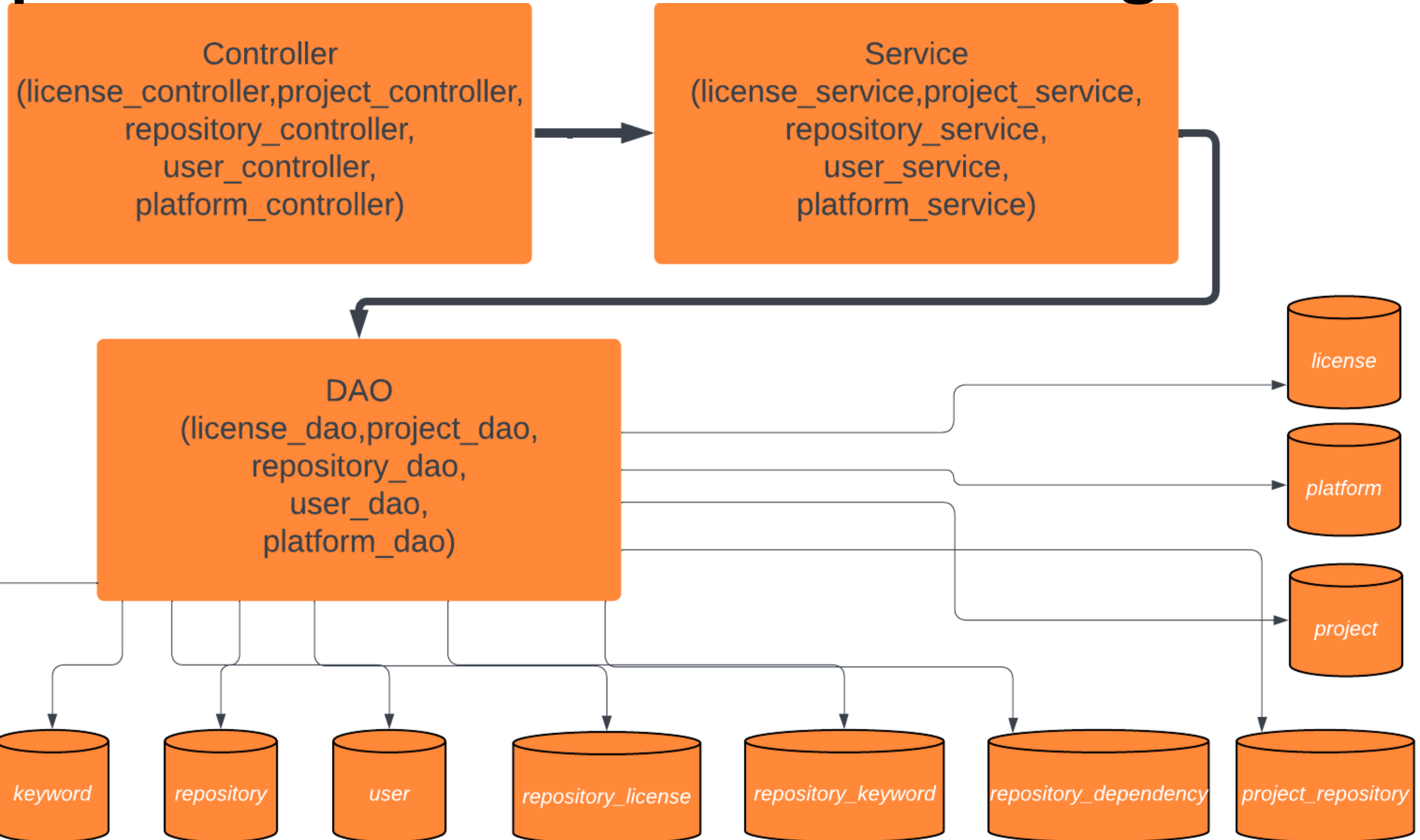
# API's Libraries.io

The Libraries.io API allows you to search for open source components and retrieve information about them. Here is a breakdown of how to use the Libraries.io API:-

- Registration: Create an account and obtain an API key from Libraries.io. The API key will be used to authenticate your requests.
- Search for components: Use the search endpoint to find open source components based on various parameters such as keywords, licenses, languages and platform.
- Retrieve required information: Once we obtain the required component, we make a request to the API endpoint for retrieving various components information. This may include details like the component's name, description, version, license, repository URL and maintainers.



# Implementation of backend API using Flask



# Elastic Search - Fuzzy Logic

Fuzzy logic is a technique used in information retrieval to account for approximate matches and handle typographical errors or variations in search queries.

- Fuzzy Query: Elasticsearch provides a Fuzzy Query feature that allows you to search for terms that are similar to a specified term, taking into account variations like misspellings user can search for the components even if the keyword is incorrect.
- Edit Distance: Fuzzy matching in Elasticsearch is based on the concept of edit distance. Elasticsearch uses the Levenshtein distance algorithm to calculate the edit distance.



# Demo

## Open Source Finder

apache

Q

John Doe

**org.apache.spark:spark-sql-kafka-0-10\_2.12**

The Apache Software Foundation provides support for the Apache community of open-source software projects. The Apache projects are characterized by a collaborative, consensus based development proces...

Latest Stable Release 3.2.4 - Updated 10/04/2023

**org.apache.flink:flink-sql-parser**

The Apache Software Foundation provides support for the Apache community of open-source software projects. The Apache projects are characterized by a collaborative, consensus based development proces...

Latest Stable Release 1.17.0 - Updated 17/03/2023

**org.apache.spark:spark-sql-kafka-0-10\_2.11**

The Apache Software Foundation provides support for the Apache community of open-source software projects. The Apache projects are characterized by a collaborative, consensus based development proces...

**Platforms**

☐ NPM

☐ Maven

☐ Go

☐ Pypi

☐ NuGet

☐ Packagist

☐ Rubygems

☐ Cargo

☐ CocoaPods

☐ Bower

**Licenses**

☐ BSD-3-Clause

☐ MPL-2.0

☐ MIT

☐ Apache-2.0

☐ GPL-3.0

# Open Source Finder

apache



John Doe



## My Projects

hpe

hpe2

new project

hpe3

new

new1

new3

## Repositories

sql

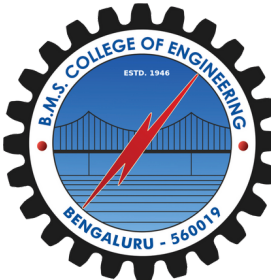
mongo-sql

org.apache.flink:flink-sql-parser

react-bootstrap



**Hewlett Packard**  
Enterprise



## org.apache.flink:flink-sql-parser

The Apache Software Foundation provides support for the Apache community of open-source software projects. The Apache projects are characterized by a collaborative, consensus based development process, an open and pragmatic software license, and a desire to create high quality software that leads the way in its field. We consider ourselves not simply a group of projects sharing a server, but rather a community of developers and users.

Latest Release - 1.17.0 Updated on 17/03/2023

[Home Page](#)[Download](#)

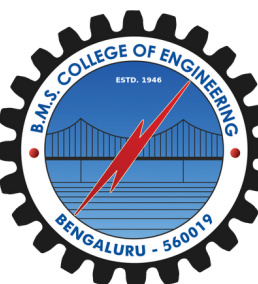
### Rank 17

Stars	21.1K
Forks	11.9K



- ☒ hpe
- ☐ hpe2
- ☐ new project
- ☒ hpe3
- ☐ new
- ☐ new1
- ☒ new3
- + New Project

**Hewlett Packard**  
Enterprise

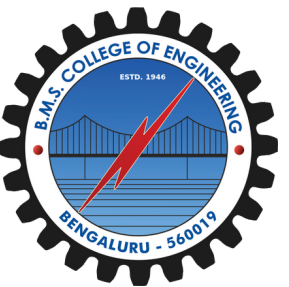


# Learnings

- Introduction to fuzzy logic using Elastic-search.
- Good understanding of mvc architecture.
- Introduction to openai and langchain to train a large language model for building a generative AI.



**Hewlett Packard**  
Enterprise

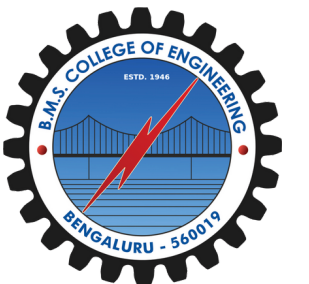


# Next Step

- Build a generative AI chatbot by training a large language model with data extracted from the database.
- Integrate a user-friendly form that enables users to manually input and add open-source components to the system.
- Implement a validation mechanism to ensure the integrity and quality of the fetched open-source components.



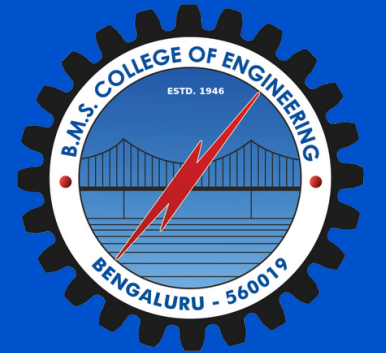
**Hewlett Packard**  
Enterprise



# Any Questions?



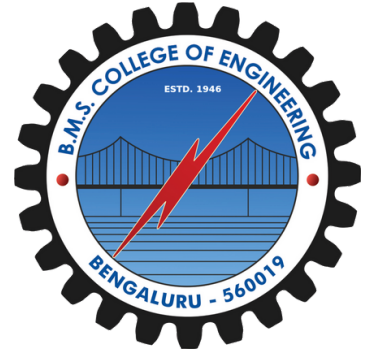
**Hewlett Packard**  
Enterprise







**Hewlett Packard**  
Enterprise



# Thank You