

# **[PRD] Basic Search App**

# **Summary:** Build a basic search application that can search the contents of documents stored in cloud storage like Google drive, Dropbox, Box etc

|  |  |
| --- | --- |
| **Created:** May 18, 2021  **Owner:** gourishhebar@gmail.com  **Contributors:** gourishhebbar@gmail.com | **Status:**  **WIP** |



## Problem

When a user has a vast number of documents stored in the cloud, reading or extracting specific useful information from the documents is difficult.The user will have to open every document and go through it to understand and decide whether the information that it contains is relevant.

## Personas

* **Affected Persona 1** Person who has a vast number of documents and can't retrieve relevant information.

## Requirements and Phases

|  |  |
| --- | --- |
|  | **Requirements** |
| **Phase 1**: Module to retrieve documents from Google Drive | With a Google Drive link, PDF and Docx files should be retrieved. |
|  |
| **Phase 2:** Module to read contents from a file | Read the contents from PDF and Docx files. |
| **Phase 3:** Module to populate ElasticSearch database | Feed the contents read by PDF and Docx file to ElasticSearch database |
| **Phase 4 :** REST API development | The exposed REST API retrieves documents based on the search term/token |
| **Phase 5 :** Simple front-end page | Build a simple front end app to communicate with the search application and display the results. |

### **Phase 1**: Module to retrieve documents from Google Drive

Requirement:

A google drive storage must contain files. Files of the type PDF and Docx need to be retrieved.

Acceptance Criteria:

* **All the documents** must be retrieved from the cloud, with the Google Drive link.

### **Phase 2**: Module to read contents from a file

Requirement:

Read all the contents from each of the PDF and Docx files, by using the API provided by Google Drive

Acceptance Criteria:

* **All the contents** of each file needs to be read.

### **Phase 3**: Module to populate ElasticSearch database

Requirement:

Feed the contents read by PDF and Docx file and index them using the ElasticSearch database service.

Acceptance Criteria:

* **The files** retrieved should be indexed.

### **Phase 4**: REST API development

Requirement:

The exposed REST API retrieves documents based on the search term/token.

Acceptance Criteria:

* **A working REST API** which takes the search term as an input and returns a list of files and their HTTP URLs that contain the term.

### 

### **Phase 5**: Simple front-end page

Requirement:

Build a simple front end app which acts as an interface to the user to enable him to communicate with the search application and get the results.

Acceptance Criteria:

1. **A front-end page** that is responsive and displays the retrieved list of files based on the search term.

## User Research

NA

## Approvals

NA