Product Requirements Document

Title	*Search Files*
Owner	*Swapnil Darji*
Contributors	*Swapnil Darji*
Resources	*https://github.com/swapnildarji/search-files*
Status	*Beta version Launched*
Last updated	*15th Dec 2024*

Problem

In today's digital landscape, many organizations and individuals store a vast amount of files in cloud storage services like Amazon S3. However, as the volume of documents grows, manually searching through them to locate specific information becomes an increasingly difficult and time-consuming task. Traditional cloud storage systems often provide basic file organization features but lack robust search functionality that can scan the contents of the documents. This inefficiency leads to frustration, delays, and decreased productivity for users who need to quickly access the right information from within a sea of documents.

Impact

- **Time Wasted***: Users spend too much time searching for specific content in files.*
- **Increased Frustration**: Manual file searching leads to frustration and inefficiency.
- **Incomplete Search Results**: Basic search features often miss content within documents.
- **Slower Collaboration**: Difficulty finding relevant documents slows down team collaboration.
- **Scaling Issues**: As the volume of files grows, searching becomes increasingly difficult.

Solution Outline

> ***Search-files*** *is a cloud-based application designed to help users efficiently search and retrieve documents stored in cloud storage services (e.g., Amazon S3). The system enables full-text search across a variety of document types (.txt, .pdf, .docx) and allows users to upload new documents for indexing and later retrieval.*

Key features

- ***Provide a seamless search experience*** *for users to find documents based on content stored in cloud storage (AWS S3).*
- **Allow file uploads** with support for `.txt`, `.pdf`, and `.docx` formats, and automatically index content for searchability.
- **Enable full-text search** capabilities using AWS OpenSearch, allowing users to search by keywords and retrieve matching documents.
- **Ensure ease of use and intuitive interface** with React.js on the frontend and Node.js on the backend, providing a responsive, user-friendly experience.
- **Provide cloud scalability** with serverless deployment on Vercel, ensuring smooth scaling during high traffic periods.

User Stories

User Story 1: Search for Documents

As as user, I want to search for specific content inside the documents stored in the cloud so that I can easily find the information I need.

Acceptance Criteria:

- Users can enter a search query in a search bar.
- The system returns a list of documents containing the search term, including metadata (e.g., document name, date uploaded).
- The search results should display a snippet of content from the document where the search term is found.

User Story 2: Upload Documents

As a user, I want to upload documents to the cloud storage so that I can store and search for them later.

Acceptance Criteria:

- Users can upload `.txt`, `.pdf`, and `.docx` files.
- The maximum file size for uploads is limited to 5MB.
- The system processes the file upon upload, indexing its content for future searches.

Milestones

Target date	Milestone	Description	Exit criteria
15-12-2024	*Internal pilot*	*Internal testing*	*No P0 or P1 bugs.*
15-12-2024	*Launch 1 (Beta)*	*Available to all users*	*Basic Search Functionality*
23-12-2024	*Launch 2*	*Available to all users*	*Search Functionality with sorting and filtering*
30-12-2024	*Launch 3*	*Available to all users*	*File uploading Functionality*

Appendix

Considerations

Files uploaded to AWS S3 should be secured using a signed URL. The URL should expire X minutes after it is generated.

Changelog

Date	Change
15-12-2024	Launched Initial Version of Application