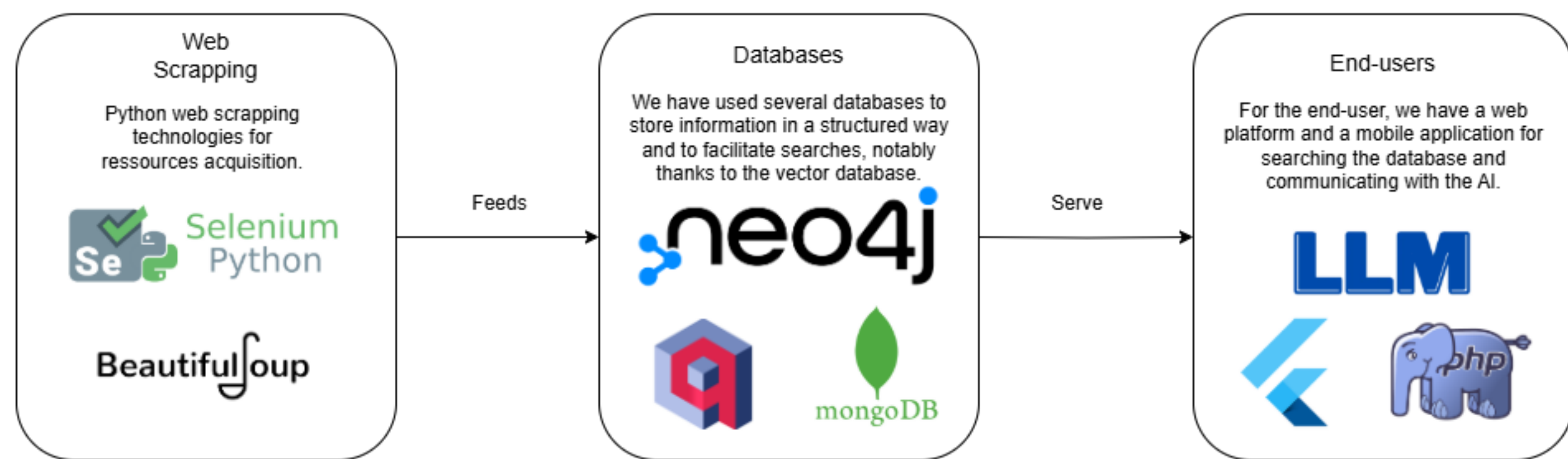


Abstract

Our project aims to facilitate the work of professionals teaching construction trades by providing them with a platform that centralizes information, allows for searches, and enables communication with a specialized AI. To achieve this, we have used web scraping technologies, new types of databases, and artificial intelligence.

Overview



Why BuildAI is a Game-Changer?

BuildAI is an innovative platform that **revolutionizes access to construction-related knowledge** by leveraging AI-driven data retrieval and structuring.

- **Centralized Knowledge:** Aggregates diverse construction resources into a single accessible hub.
- **AI-Powered Search:** Utilizes **semantic search and LLMs** to retrieve the most relevant information.
- **Hybrid Database System:** Efficiently manages data using **vector, graph, and document databases**.
- **Real-Time Data Extraction:** Continuously scrapes and updates industry-related content.
- **Educator-Friendly:** Tailored for instructors to easily **build and customize training materials**.

Key Impacts of BuildAI

BuildAI transforms the way professionals in the construction sector access and utilize knowledge. By combining AI-driven search, structured databases, and real-time data updates, it offers significant benefits:

- **Faster Information Retrieval:** Eliminates manual searching by structuring vast industry knowledge into an easily accessible system.
- **Smarter Learning Experience:** Uses AI-powered recommendations to provide tailored educational content for instructors and trainees.
- **Seamless Access Across Devices:** Available as a **web platform and mobile application**, ensuring knowledge is at your fingertips anytime, anywhere.
- **Interconnected Knowledge Base:** Integrates **vector, graph, and document databases** to enhance contextual understanding and resource linking.
- **Scalable and Future-Ready:** Designed to **expand beyond construction** into other industries requiring structured knowledge management.

How Web Scraping Works

Web scraping is the automated process of extracting structured information from websites. BuildAI leverages web scraping to collect and organize construction-related data from multiple online sources. The process follows these key steps:

1. **Identifying Target Websites:** Selection of relevant construction-related sources, such as regulatory documents, training materials, and technical blogs.
2. **Sending HTTP Requests:** Tools like Selenium and BeautifulSoup are used to access web pages and retrieve their HTML content.
3. **Parsing and Extracting Data:** The raw HTML is analyzed to extract structured information, such as text, tables, images, or metadata.
4. **Cleaning and Formatting:** The extracted data is processed to remove unnecessary elements, correct inconsistencies, and standardize the format.
5. **Storing in Databases:** The cleaned data is stored in **vector, graph, and document databases** for efficient retrieval and indexing.
6. **Continuous Updates:** The system periodically scrapes updated content to ensure that professionals always access the latest industry knowledge.

Web scraping allows BuildAI to automate data collection at scale, ensuring that construction professionals and educators have access to **comprehensive, structured, and up-to-date resources**.

Database Architecture: Neo4j, MongoDB, and Qdrant

BuildAI leverages a hybrid database architecture to efficiently manage and retrieve construction-related data. Three specialized databases are used to handle different types of information:

Neo4j - Graph Database

- Designed to store **relationships** between entities such as materials, techniques, and regulations.
- Enables **networked queries**, making it ideal for recommendation systems and knowledge graphs.
- Provides fast **pathfinding algorithms** to discover connections between concepts.

MongoDB - Document Database

- Stores **unstructured and semi-structured** data such as training documents, reports, and manuals.
- Uses **JSON-like documents**, allowing flexible and scalable data storage.
- Ideal for **storing diverse content types** in an intuitive, hierarchical structure.

Qdrant - Vector Database

- Optimized for **high-dimensional vector search** and similarity matching.
- Enables **semantic search** using AI-generated embeddings from construction-related texts.
- Facilitates fast **retrieval of relevant information** based on contextual meaning.

By combining these three databases, BuildAI ensures **efficient knowledge representation, fast queries, and scalable data storage**, making it a robust solution for professionals in the construction industry.

BuildAI: Web Platform, Mobile App, LLM, and Search Engine

BuildAI provides a multi-platform solution to ensure seamless access to construction-related knowledge. The system consists of:

Mobile Application (Flutter)

The BuildAI mobile app, developed in **Flutter**, offers:

- A responsive and intuitive interface for on-the-go learning.
- Seamless integration with the AI-powered search engine.
- Offline access to stored construction documents and resources.

Web Platform (PHP)

The web version of BuildAI is built using **PHP** and designed to:

- Provide a full-featured experience for managing and browsing resources.
- Allow instructors to upload and structure educational materials.
- Connect with Neo4j, MongoDB, and Qdrant for efficient data retrieval.

Large Language Model (LLM)

BuildAI integrates a **custom-trained LLM** to enhance information retrieval:

- Processes natural language queries for intelligent responses.
- Generates summaries of construction documents for quick insights.
- Provides contextual assistance to users based on their search intent.

AI-Powered Search Engine

The **search engine** combines multiple AI-driven techniques:

- **Vector search (Qdrant):** Finds relevant documents based on meaning rather than keywords.
- **Graph-based retrieval (Neo4j):** Explores relationships between concepts.
- **Document indexing (MongoDB):** Allows structured search through large datasets.

By combining these technologies, BuildAI delivers an **integrated, scalable, and user-friendly** platform for construction professionals and educators.

Project Architecture

