

Florida International University
Knight Foundation School of Computing and Information Sciences

Software Engineering Focus

Final Deliverable

Project Title: Material Atlas for Coastal Construction

Team Members: Kyle Grabova, Cesar Barrios, Giovanni Ruiz, Sejal Nathu-Hari, Chris Daniel Lantigua

Product Owner(s): Lyla Wu

Mentor(s): Lyla Wu

Instructor: Masoud Sadjadi

The MIT License (MIT)

Copyright (c) 2016 *Florida International University*

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

Abstract

This document outlines the development and functionalities of Material Atlas for Coastal Construction, a platform aimed at transforming the sharing and visualization of construction material data tailored for coastal environments.

Developed using Next.js and React for the front end, and .NET Core with C# for user management, the platform boasts a secure and intuitive interface. Central to its architecture is a Neo4j graph database, which facilitates seamless navigation and exploration of interconnected data sourced primarily from research papers.

A key highlight of this platform is the integration of the OpenAI API with Langchain, leveraging GPT's natural language processing capabilities to automate the extraction of data from material science research paper PDFs.

This innovative approach significantly streamlines the process of populating the database, enhancing the efficiency of data assimilation for research purposes. Designed to cater to professionals in fields such as engineers, material scientists, and architects, the atlas serves as an indispensable resource for accessing, comparing, and utilizing material science data to promote sustainable and resilient coastal construction practices. Through its user-friendly interface and advanced data processing capabilities, Material Atlas aims to empower experts across disciplines to make informed decisions and drive positive change in coastal development projects.

Table of Contents

Abstract.....	3
Table of Contents.....	4
Introduction.....	5
Current System.....	5
Purpose of New System.....	5
User Stories.....	7
Implemented User Stories.....	7
Pending User Stories.....	8
Project Plan.....	9
Hardware and Software Resources.....	9
Sprints Plan.....	10
Sprint 1.....	10
Sprint 2.....	10
Sprint 3.....	10
Sprint 4.....	10
Sprint 5.....	11
Sprint 6.....	11
System Design.....	12
Architectural Patterns.....	13
System and Subsystem Decomposition.....	13
Deployment Diagram.....	13
Design Pattern.....	13
System Validation.....	13
Glossary.....	13
Appendix.....	14
Appendix A - UML Diagrams.....	14
Appendix B - User Interface Design.....	16
Appendix C - Sprint Review Reports.....	18
Appendix D - User Manuals, Installation/Maintenance Document, Shortcomings/Wishlist Document and other documents.....	18
Shortcomings/Wishlist.....	18
References.....	18

Introduction

Welcome to the documentation for Material Atlas for Coastal Construction. This guide provides an in-depth overview of our interdisciplinary platform, designed to revolutionize the sharing and visualization of material science data tailored specifically for coastal regions. Developed with a combination of cutting-edge technologies, including Next.js, React, .NET Core with C#, and Neo4j, Material Atlas offers a secure, intuitive, and feature-rich environment for professionals across various disciplines. In this documentation, you will find detailed explanations of the platform's architecture, functionalities, and implementation. From understanding the core technologies powering Material Atlas to exploring its innovative features such as automated data extraction using the OpenAI API with Langchain, this documentation serves as your comprehensive resource for navigating and utilizing the platform effectively.

Whether you are a researcher seeking to access and analyze material data for coastal construction projects or an engineer/architect looking to make informed decisions regarding sustainable and resilient construction practices, this documentation equips you with the knowledge and tools necessary to leverage Material Atlas to its fullest potential.

Current System

In the realm of coastal construction and environmental sustainability, the current system presents significant challenges. With the onset of climate change, rising sea levels, and the increasing vulnerability of coastal areas, there is a pressing need for a comprehensive understanding of sustainable materials and construction techniques. However, the current landscape is fragmented, with research and knowledge scattered across various disciplines, including Marine Engineering, Material Science, Civil Engineering, and Environmental Engineering. This fragmentation leads to significant barriers in accessing, correlating, and applying crucial data effectively in coastal construction projects. The absence of a centralized repository or platform for sharing and analyzing information about construction materials, their properties, applications, and environmental impacts severely limits the potential for innovative and sustainable coastal development.

The current system, with the implementation of the undermentioned changes, represents a significant evolution in the Material Atlas platform. With the integration of a fully functional OpenAI data extraction and analysis tool, users now have the capability to effortlessly upload research papers and extract relevant material data with unparalleled efficiency. This enhancement not only streamlines the process of populating the database but also ensures the accuracy and reliability of the extracted information. Moreover, the incorporation of enhanced login security features, including password hashing, reinforces the platform's commitment to protecting user data and maintaining privacy. Alongside these advancements, significant improvements have been made to the user interface, enhancing usability and aesthetics to provide a seamless and intuitive experience for users navigating the platform. These collective changes have elevated Material Atlas into a robust and user-friendly resource, empowering professionals across disciplines to access, compare, and utilize material data with unprecedented ease and efficiency.

Purpose of New System

The primary purpose of this system is to bridge the gap between various fields of expertise involved in coastal construction and environmental sustainability. By providing a centralized hub for knowledge sharing and collaboration, the platform aims to facilitate the exchange of information among experts in Engineering, Material Science, Architecture Design, Ecology, and more. The system leverages advanced technologies, including Next.js with React for the front end, .NET Core with C# for robust user management, and a combination of PostgreSQL and Neo4j graph databases for managing user data and research content. A significant innovation of the platform is the utilization of the OpenAI API with Langchain for automated data extraction from research papers, streamlining the incorporation of vast amounts of information into the Neo4j database. This feature not only enhances the efficiency of data integration but also supports the visualization of complex relationships between materials and their applications. The Material Atlas aims to be a pioneering tool in the field, offering graphical representations and comparisons of materials, ultimately facilitating more informed, sustainable, and resilient coastal construction practices.

User Stories

- Research AI tools for extracting formulas/methods from research papers.
- Set up a local instance of the previous website.
- Gather project requirements from Product Owner.
- Familiarize the team with the chosen tech stack.
- Develop a strategic plan for project execution.

Implemented User Stories

- Develop and refine a data extraction tool.
- Test the data extraction tool with research papers.
- Fix login and sign-up features.
- Implement encryption for security.
- Finished creating an API to interact between the data extraction script and ChatGPT API.
- Started and finished creating an Admin Dashboard for editing and tracking traffic on the website
- Successfully integrated LangChain dependency to enhance the extraction of data from PDFs.
- Continue to improve the data extraction algorithm through refinement and optimization in order to maintain a high level of accuracy.
- Testing of diverse research papers for comprehensive testing of the tool for data extraction to ensure it achieves robust performance.
- Make sure to fix identified bugs with the login and sign-up features, thus improving user experience.
- Configured and implemented encryption protocols in the login database in order to strengthen security measures.
- Complete work on the development and deployment of Admin Dashboard for efficient content management and traffic monitoring.
- Test, validate, and make sure the accuracy of material and property identification is on point with ChatGPT API.
- Contribute in research of image analysis using ChatGPT API and OpenCV to translate graph interpretation.
- Keep refining ChatGPT prompts to improve its accuracy and relevance.
- Always keep looking at efficiency and reliability of API integration between data extraction and ChatGP

Pending User Stories

- Implementing Dashboard Admin into live website
- Formatting response output for ChatGPT API

Project Plan

The project was designed and implemented with the maximum level of care, as agile development methodologies were used for flexibility and efficiency. This chapter therefore contains details on how the project was planned and implemented, incorporating rigidity and flexibility to cover all aspects of the project. Sprint planning was very key as a parameter on how to go about development and being able to make changes or adjustments where necessary. Every sprint was planned to the last detail and executed accordingly. Also in this section, how the chosen project components were decided is mentioned, including software and hardware. Each component was chosen because of its compatibility, functionality, and support for the project's goals. This predefined the solid base of the project team.

Hardware and Software Resources

The following is a list of all hardware and software resources that were used in this project:

Hardware:

- Not applicable
- Online clouding

Every member of the group used their personal laptops, all of them were Macbooks pro.

Software:

- Neo4j
- Python
- JavaScript
- React
- Visual Studio Code
- Next js
- OpenAI API

Sprints Plan

Sprint 1

- User Story 1: Research AI tools to extract formulas / methods from research papers
- User Story 2: Work with team to get instance of previous website running locally
- User Story 3: Gather information from PO to understand the project requirements
- User Story 4: Research and familiarize team with tech stack
- User Story 5: Create a strategic plan of the work to be carried out
- User Story 6: Discuss plan with PO

Sprint 2

- User Story 1: Work on first iteration of data extraction tool
- User Story 2: Test data extraction tool with 5 research papers
- User Story 3: Implement improvements for data extraction tool based on data collected from testing with material science research paper
- User Story 4: Fine tune data extraction tool parameters
- User Story 5: Research bugs within login and sign up processes
- User Story 6: Continue researching AI models for data extraction

Sprint 3

- User Story 1: Fixing Log in and Sign in features
- User Story 2: Implementing encryption for security protocols
- User Story 3: Continue fine tuning data extraction and plan a method to sort them for easy access
- User Story 4: Instantiating and updating previous version of database
- User Story 5: Researching methods to extract and analyze images from research papers
- User Story 6: First stages of creating an API to interact with between the data extraction script and ChatGPT API

Sprint 4

- User Story 1: Testing LangChain dependency and linking it to the data extraction script
- User Story 2: Testing ChatGPT accuracy to determine materials and their properties
- User Story 3: Adjusting and rewriting prompts to determine the most accurate responses
- User Story 4: Working on a way to implement the script into the existing website
- User Story 5: Still researching ways to analyze graph images using ChatGPT API and OpenCV
- User Story 6: Working on Admin Dashboard requested by PO

Sprint 5

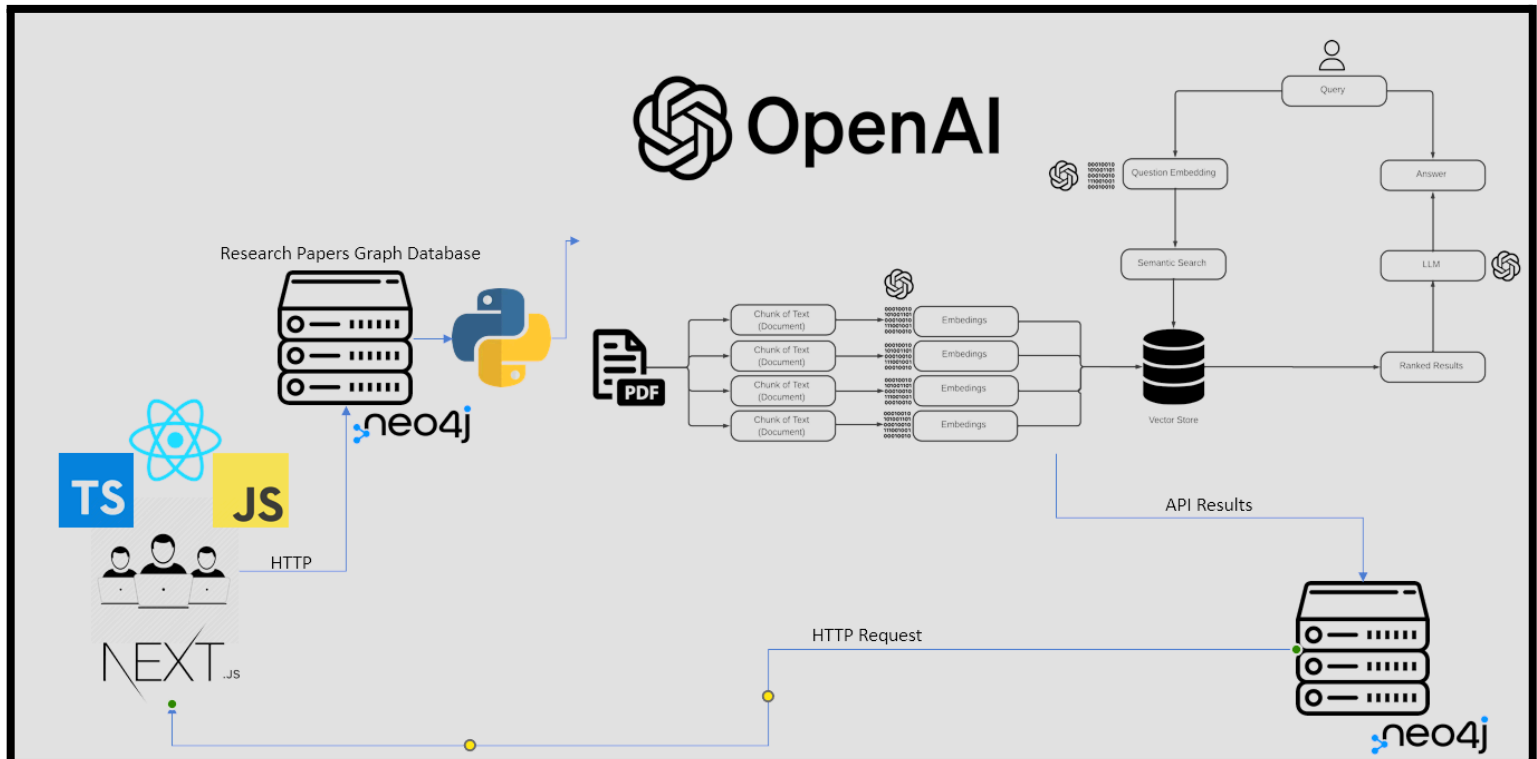
- User Story 1: Making progress on Admin Dashboard
- User Story 2: Full implementation of ChatGPT with the data extraction script
- User Story 3: Fix minor bugs related to the data extraction, admin dashboard and log in feature
- User Story 4: Testing ChatGPT API's accuracy and consistency by frequent communication with an expert on the topic (PO)
- User Story 5: Researching way to analyze graph images across different research papers

Sprint 6

- User Story 1: Finishing up Admin Dashboard
- User Story 2: Implementing working solution of data extraction into the website
- User Story 3: Troubleshooting and fixing remaining issues regarding the Admin Dashboard and data extraction implementation
- User Story 4: Still researching way to analyze graph images across different research papers
- User Story 5: Continue testing ChatGPT API's accuracy and consistency.

System Design

Upon receiving submission of a PDF document, the backend system basically extracts and segregates the information contained therein into components. After this, such components are processed through the ChatGPT API as embeddings, with a given prompt to provoke responses. This response, crafted with information such as material names, properties, etc., derived from the research paper PDF submitted.



Architectural Patterns

System and Subsystem Decomposition

Deployment Diagram

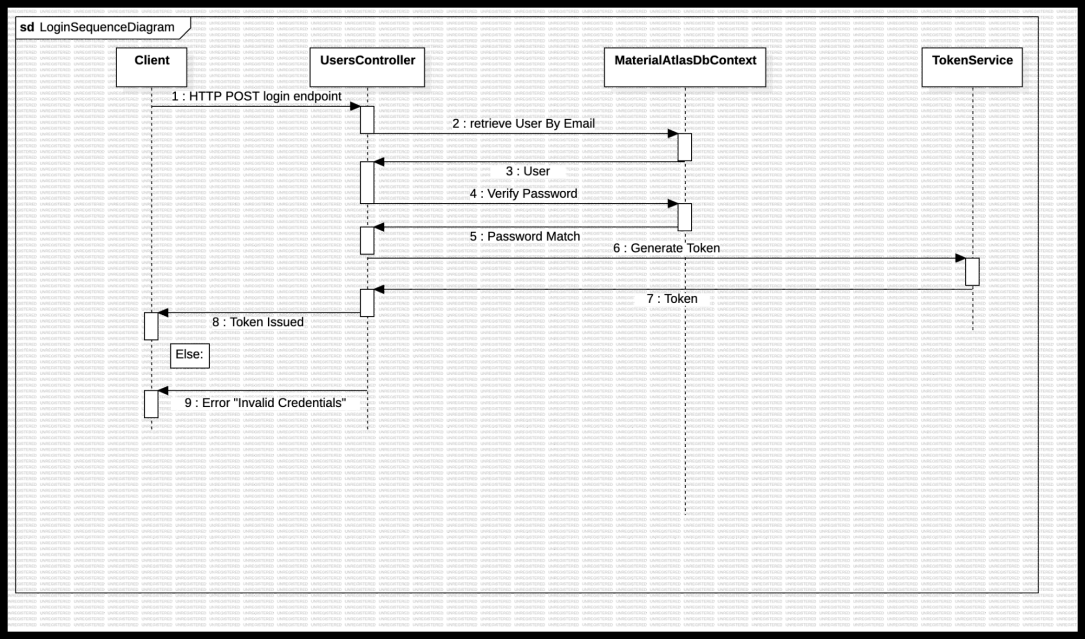
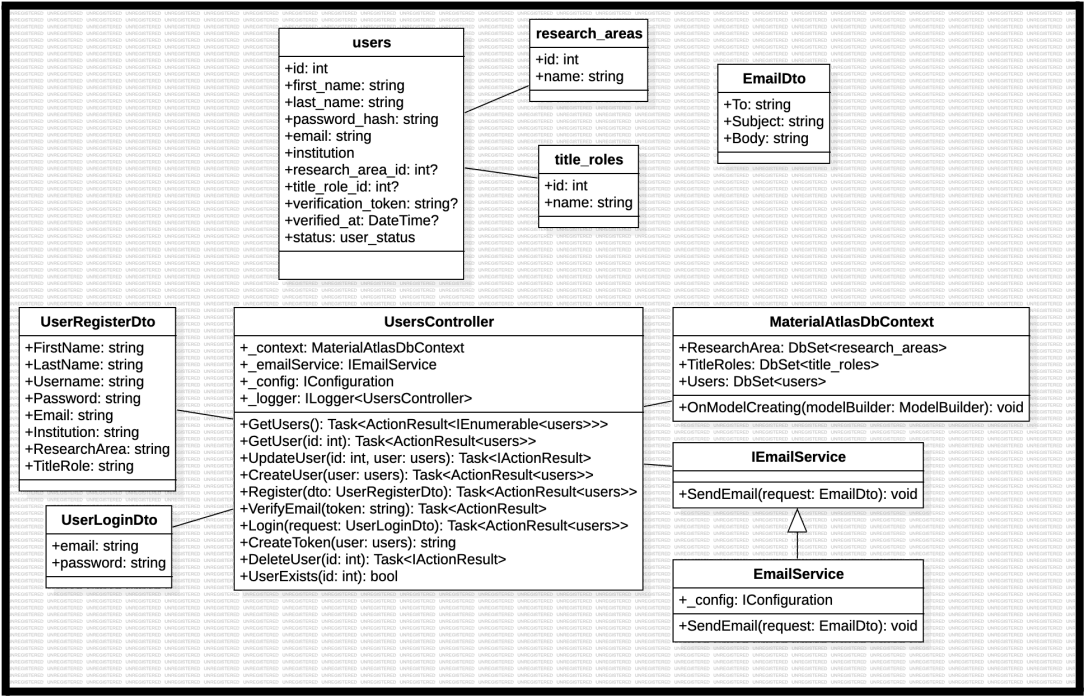
Design Pattern

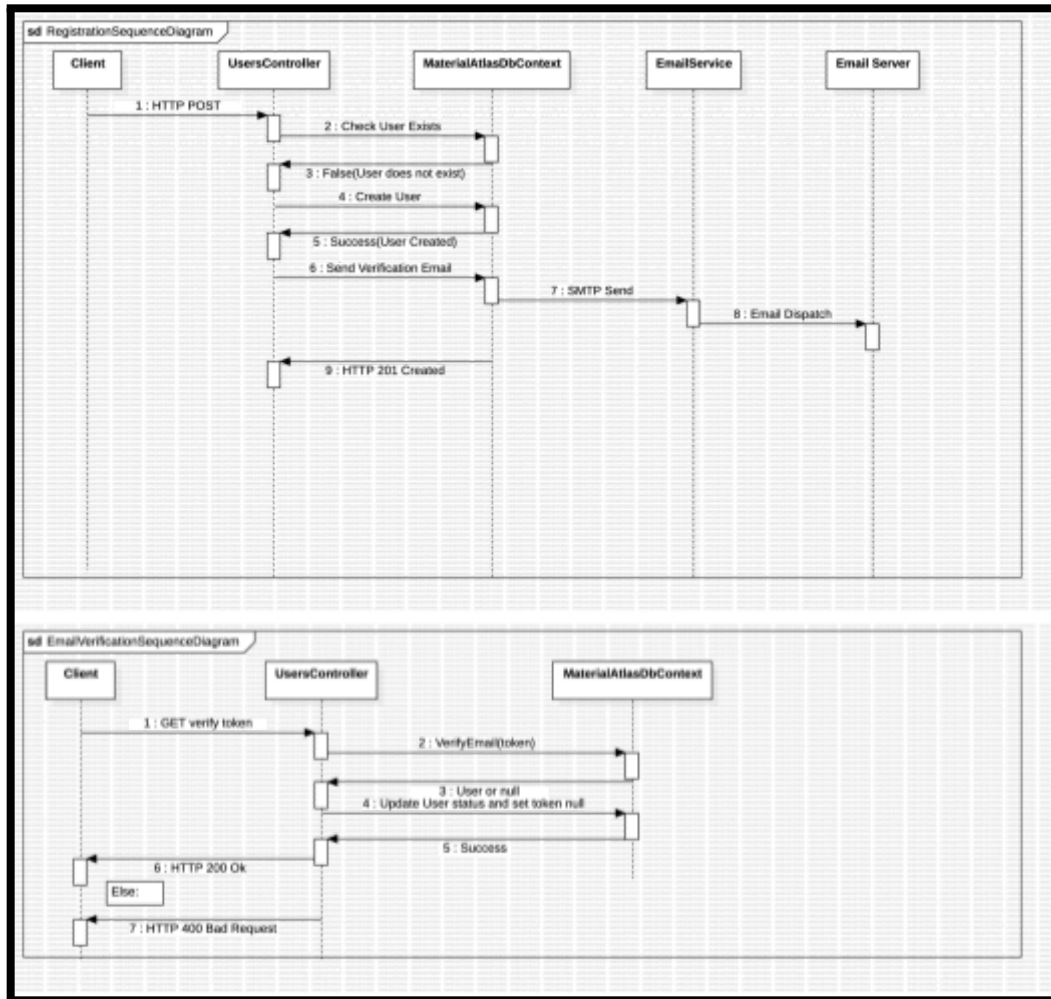
System Validation

Glossary

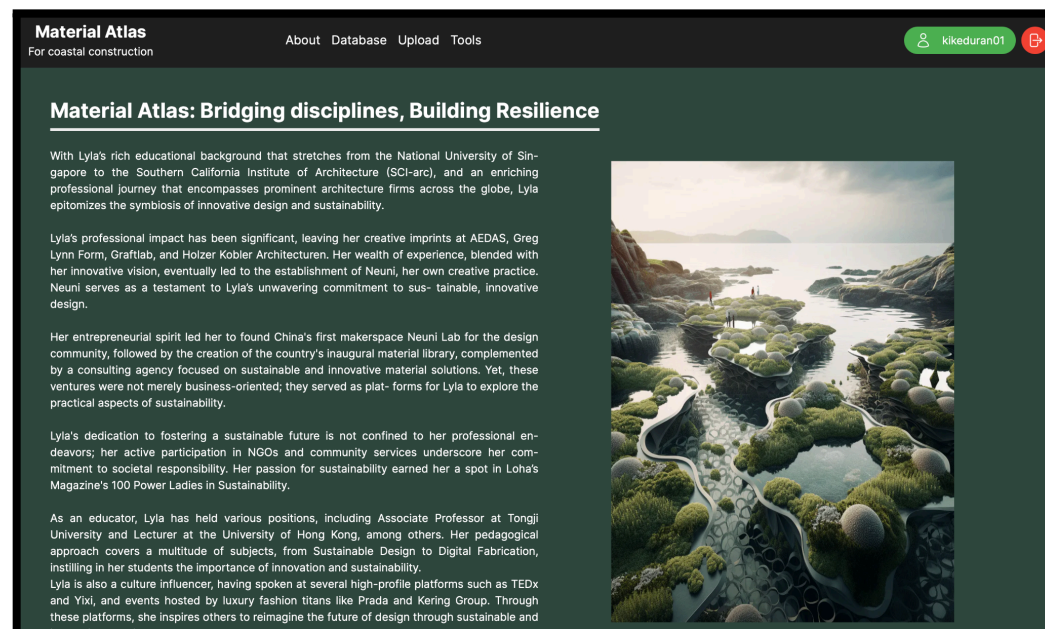
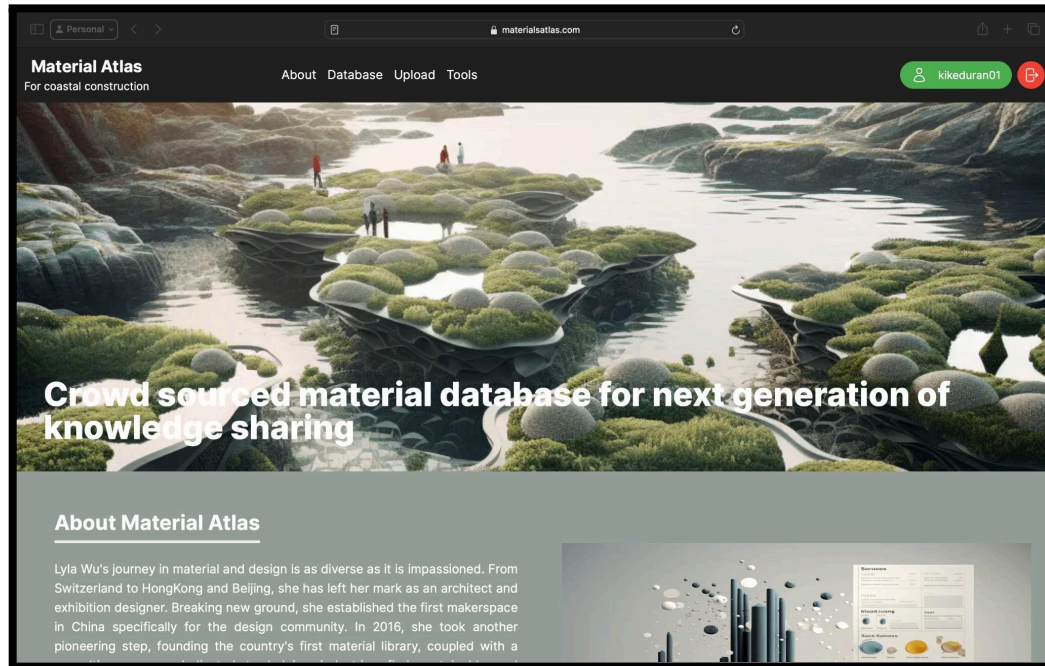
Appendix

Appendix A - UML Diagrams





Appendix B - User Interface Design



The image shows a web browser window displaying the Material Atlas website. The browser's address bar shows the URL `materialatlas.com`. The page has a dark header with the text "Material Atlas For coastal construction" on the left, navigation links "About Database Upload Tools" in the center, and "Sign In" and "Register" buttons on the right. The main content area is white and features a "Register" form. The form includes fields for "First Name" (with "John" entered), "Last Name" (with "Doe" entered), "User Name", "Email", "Password", "Institution", "Research Area" (a dropdown menu with "Marine Science" selected), and "Role and Position" (a dropdown menu with "Professor" selected). A "Register" button is at the bottom of the form. Above the form, there is a dark overlay with a "Sign in with Credentials" button, which has "Username" and "Password" input fields above it.

Material Atlas
For coastal construction

About Database Upload Tools

Sign In Register

Register

First Name Last Name

John Doe

User Name

Email

Password

Institution

Research Area

Marine Science

Role and Position

Professor

Register

Appendix C - Sprint Review Reports

Appendix D - User Manuals, Installation/Maintenance Document, Shortcomings/Wishlist Document and other documents

Shortcomings/Wishlist

- Integrating the AI model into the existing website
- Further enhancing the UI of the Material Atlas website
- Implementing error handling for the AI responses
- Enhancing search functionality to provide more precise results
- Expanding the database to include a wider range of materials and properties
- Implementing user feedback mechanisms to continuously improve the platform
- Enhancing data visualization tools for clearer presentation of material-property relationships

References