MICHIGAN STATE UNIVERSITY

Project Plan Presentation

Semantic Search for Code and Architecture Assets

The Capstone Experience

Team Amazon

Zayd Abualfellat
Sampan Chaudhuri
Jerry Chen
Atharva Kirkole
Nicholas Li

Department of Computer Science and Engineering
Michigan State University

Spring 2025

From Students…
…to Professionals

Project Sponsor Overview

- A Big Five American IT company
- E-commerce platform
- Offers web services
  - Databases
  - LLMs
  - Hosting Platforms
  - Computing

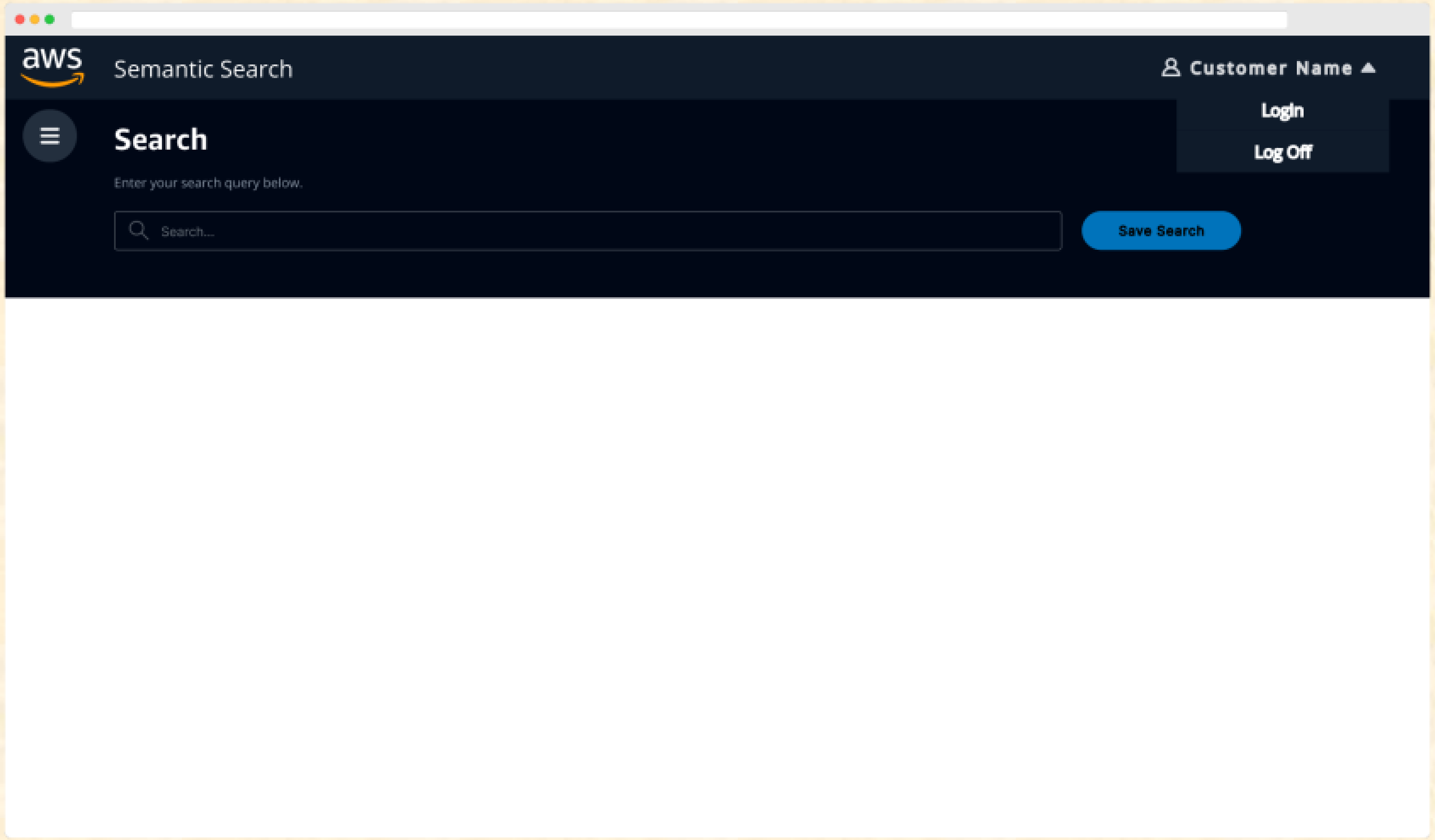
Project Functional Specifications

- Problems
  - Codebases are scattered
  - Manual code searching is inconvenient
- Solution
  - Centralize the code database
  - Search using natural language query
  - Returns relevant code results

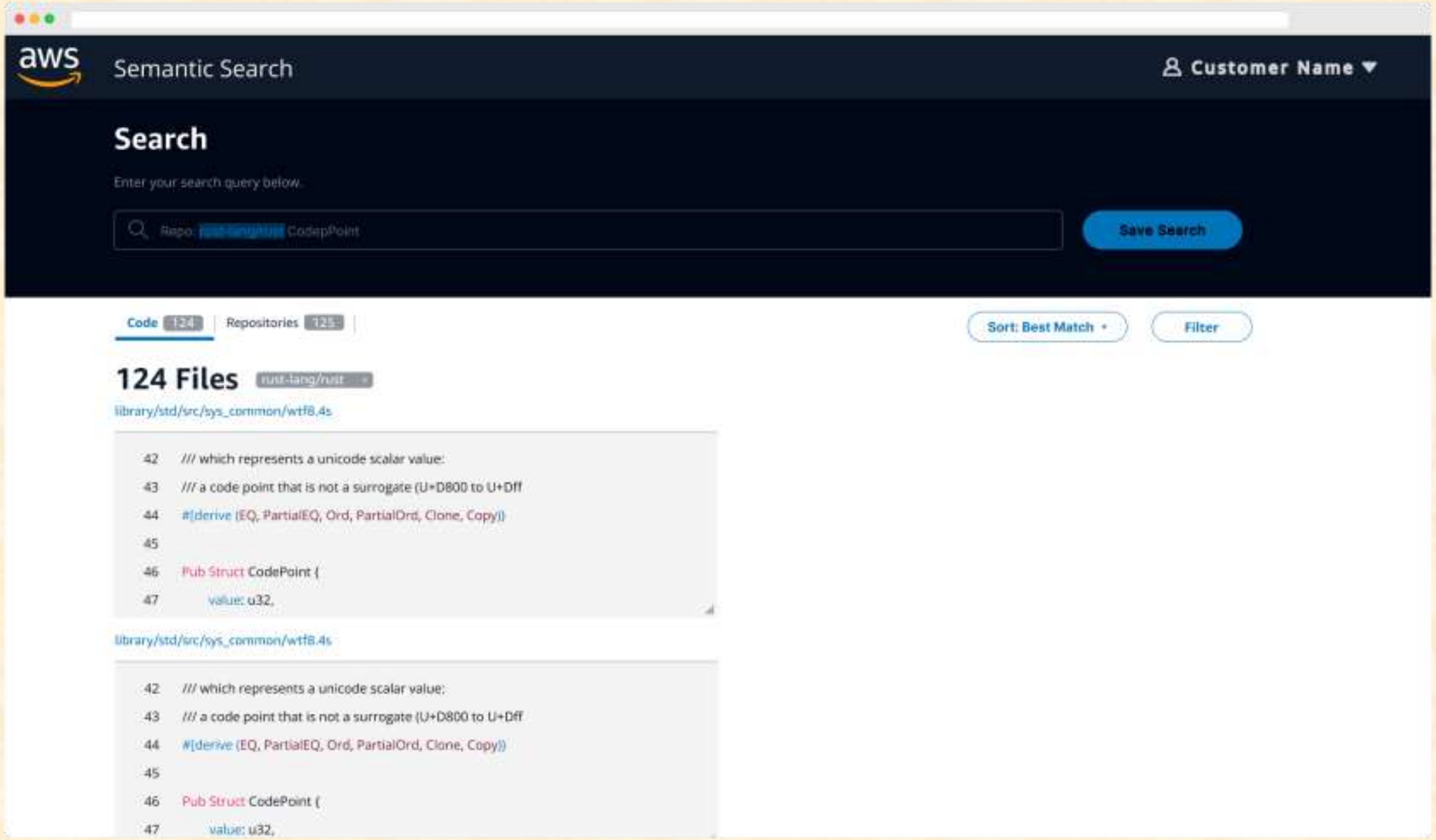
Project Design Specifications

- Intended for Amazon developers
- Enter query into search bar
- Returns top 3 relevant results
- Filter search results

Screen Mockup: Search Page



Screen Mockup: Results Displayed



Search

Enter your search query below.

Repo: rust-lang/rust CodepPoint

124 Files rust-lang/rust

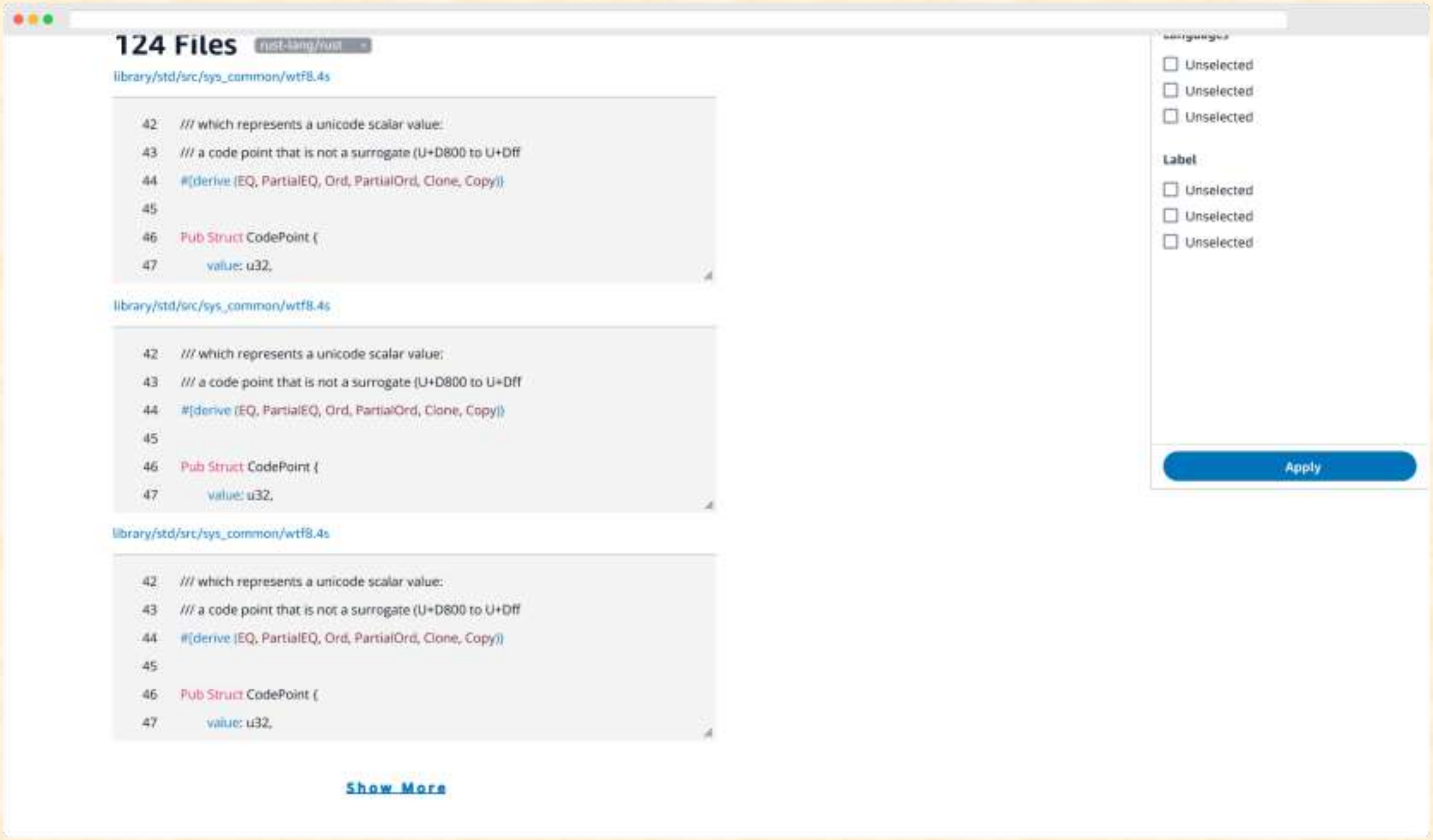
library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialDrd, Clone, Copy)]

Pub Struct CodePoint {
      value: u32,
```

library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

Pub Struct CodePoint {
      value: u32,
```

Screen Mockup: Show More Code



124 Files rust-lang/rust ×

library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

Pub Struct CodePoint {
    value: u32,
```

library/std/src/sys_common/wtf8.4s

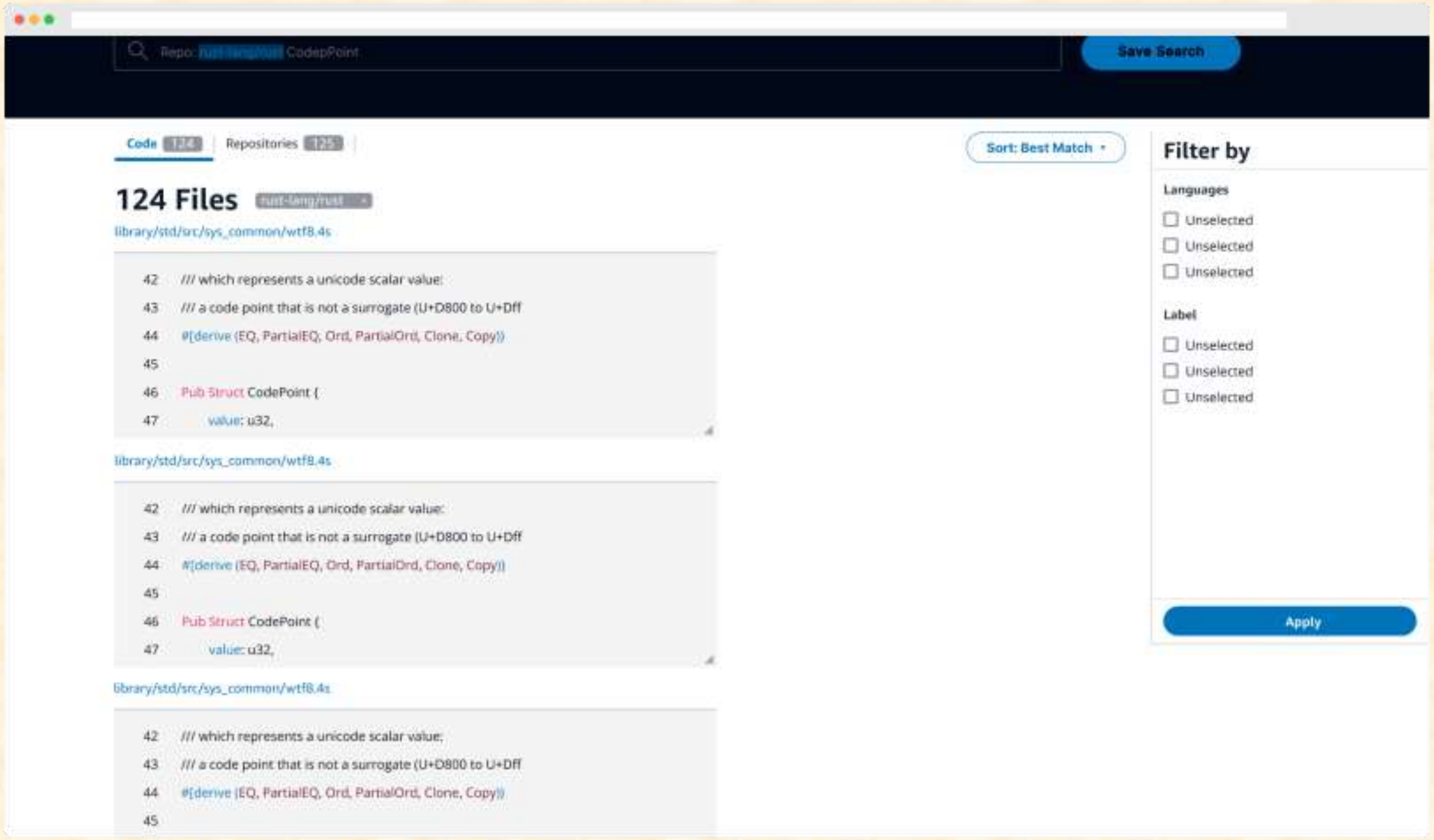
```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

Pub Struct CodePoint {
    value: u32,
```

library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

Pub Struct CodePoint {
    value: u32,
```

Show More

Screen Mockup: Filters Expanded



Repo: rust-lang/rust CodepPoint

124 Files rust-lang/rust ×

library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

Pub Struct CodePoint {
      value: u32,
```

library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

Pub Struct CodePoint {
      value: u32,
```

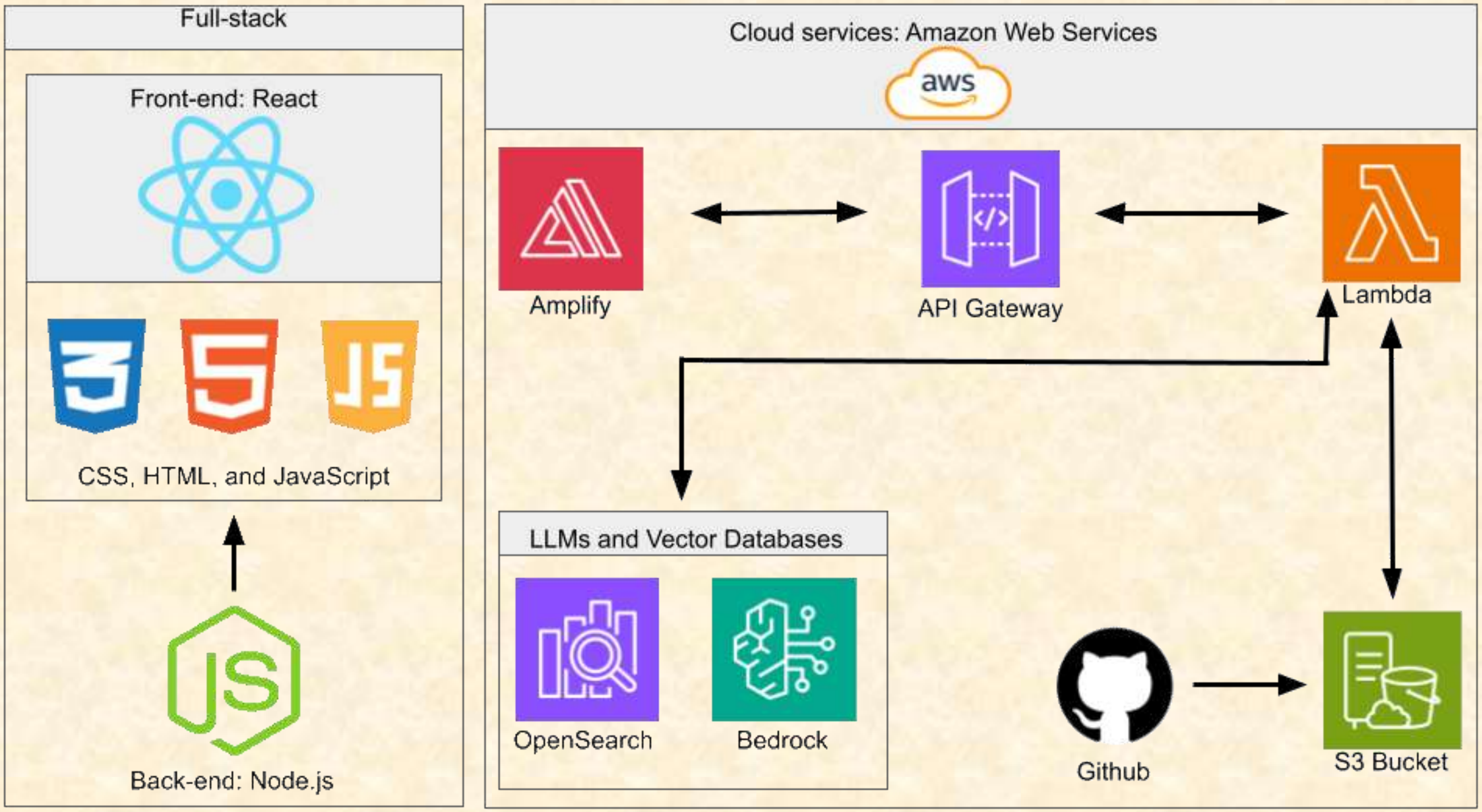
library/std/src/sys_common/wtf8.4s

```
/// which represents a unicode scalar value:
/// a code point that is not a surrogate (U+D800 to U+Dff
#[derive (EQ, PartialEQ, Ord, PartialOrd, Clone, Copy)]

```

Project Technical Specifications

- Frontend: React.js
- Backend: Node.js
- Amazon Web Services
- LLMs from Amazon Bedrock
  - Titan Text Embeddings
  - Titan Text

Project System Architecture



Project System Components

- Hardware Platforms
  - None
- Software Platforms / Technologies
  - Amazon Web Services
    - AWS Lambda
    - Amazon OpenSearch
    - Amazon Bedrock
    - Amazon S3
  - Languages
    - Python
    - JavaScript

Project Risks

- Project Expenses
  - Costs for using AWS cloud resources
  - Research and provide cost estimate to client for approval
- Rate Limiting
  - Rate limits occur with quick, consecutive calls to APIs
  - Plan to segment API calls
- Frontend and Backend Connection
  - How to send LLM-generated data from back-end to front-end
  - Implement API Gateway and Amplify solutions

Questions?

? ? ? ?

? ?

? ? ?