Release Summary
CodeCompass
By: Mann, Vincent, Evan, Nickolas, Stephen, Hao

Key User Stories and Acceptance Criteria:
- As a user, I want to have an AI-powered RAG-Chat Interface to interact with the codebase
  - User should be able to input a question about their code repository to Application
  - Application should return a relevant reply to the question to User
- As a user, I want to be able to choose an LLM model for the chatbot
  - User should be able to select the LLM model in the Application
  - User should receive responses from different LLM models from Application
- As a user, I want to have an A RAG retrieval engine retrieve relevant code chunks to the user query from the input codebase
  - User should be able to input a question to Application
  - User should receive a reply from Application
  - User should receive context on what code sections were referenced to generate the reply from Application
- As a user, I want automatic AI-generated documentation for the codebase
  - User should be able to enter documentation page
  - Application should create relevant AI-generated documentation on all files of the repository
  - User should be able to view documentation generated by the Application
- As a user, I want to be able to upload codebase from both Public & Private Github repos and Local Folders
  - User should be able to authorize codecompass on github
  - User should be able to input their github https clone link to Application
  - Application should be able to clone the link sent by user regardless of public or private repository
  - Application should index the uploaded repo or return an error on an invalid repo
  - User should be able to interact with the uploaded repo on Application using documentation, chat, and code-snippet tools
- As a user, I want to be able to upload multiple codebases (i.e., Multiple projects)
  - User should be able to create multiple projects in Application
  - User should be able to upload different repository to different projects in Application
  - User should be able to switch between codebases in Application
  - User should be able to interact with documentation, code snippet, and chat on all codebases in Application
- As a user, I want a method to display retrieved relevant code snippets from the codebase.
  - User should be able to access code-snippets page in project
  - User should be able to query Application to generate relevant code chunk based off of the codebase

- As a user, I want to have all codebases uploaded to be displayed as icons on their home page
  - User should be able to see different codebases displayed a icons on home page

Known Problems:
Backend:
- Generated Documentation and Responses may not be fully accurate or helpful
  - Hallucinations & context poisoning
- RAG Agent
  - Can sometimes choose an incorrect tool to retrieve context from.
  - Will provide the irrelevant context retrieved to the LLM without doing any reasoning/reranking or checking to see if the retrieved context is relevant to the user query
- Incorrect markdown or mermaid diagram syntax
  - Fallbacks exist, but are not comprehensive
- Lack of support for all programming languages
  - Code indexing is not language-agnostic
  - Support for indexing a programming language has to be implemented separately
  - Graph Paradigm focuses on Object-Oriented Programming Languages
    - Nodes can be classes, methods
    - Integration with Functional Programming Languages would be difficult (e.g., Haskell, Scala)
- Memory & computation limitations
  - Generating documentation & initial indexing can be slow
  - Running powerful LLMs locally requires a strong GPU
- Reindexing
  - Simply indexes the entire repository again
    - A more efficient method would be to check files that have changed (git diff) and reindex those only.

Frontend:
- Can't stop indexing once it starts, or if any errors occur during indexing.
- Choice of LLM on the frontend chat component is non-functional
  - The choice of LLM in the frontend is hardcoded to dummy api llm providers; the backend only uses local models for inference.
- The GitLab and local provider options are dummy and don't actually clone the codebase from these sources.

Product Backlog:
- As a user, I want a way to reindex projects (if changes have been committed to it) while still maintaining a way to chat with older versions of the project (i.e., not discarding older indexed projects for the new one, but rather maintaining a version history)
- As a user, I want a web-hosted instance of CodeCompass (+user guide)

- As a user, I want to be able to resize, minimize, & change placement of the different components(Chat window, Relevant Code, AI documentation) on the screen.
- As a user, I want to upload repos from GitLab, Bitbucket & Mercurial so I can interact with my code base no matter which tool is used to store it