

MICHIGAN STATE

UNIVERSITY

Project Plan Presentation

Shareholder Engagement Chatbot

The Capstone Experience

Team Ally

George Hunt
Marshal DiGiovanni
Juan Sabogal Olarte
Will Feddern
Treasure Puso

Department of Computer Science and Engineering
Michigan State University

Spring 2024



*From Students...
...to Professionals*

Project Sponsor Overview

- Financial services company
- Nation's largest all-digital bank
- HQ in Detroit, MI



Project Functional Specifications

- Context-aware chatbot
- Based on Ally's investor-focused materials
- Answer shareholder questions
- Provide source citations
- Refuse to disclose private information
- Designed to improve investor engagement

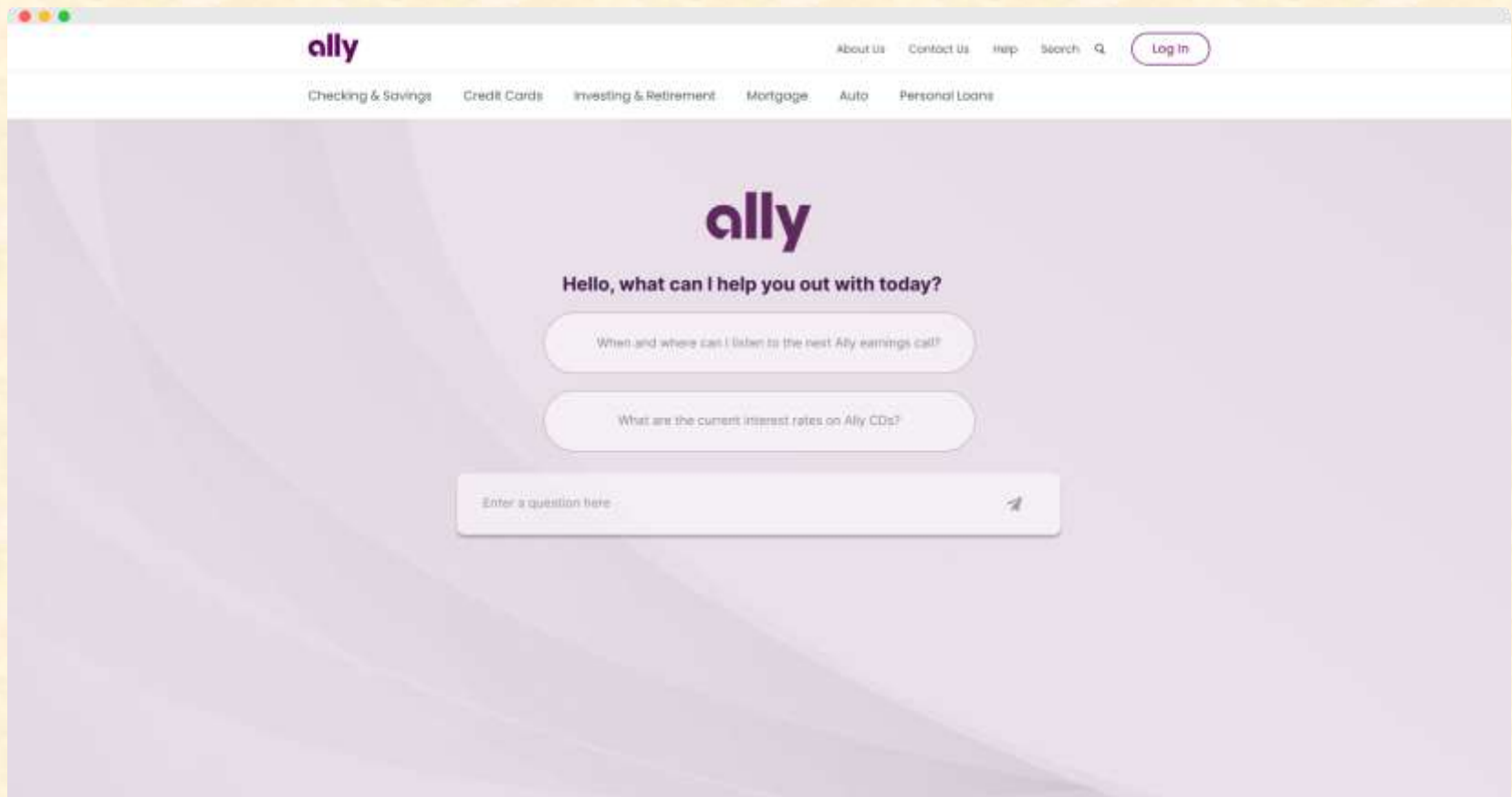


Project Design Specifications

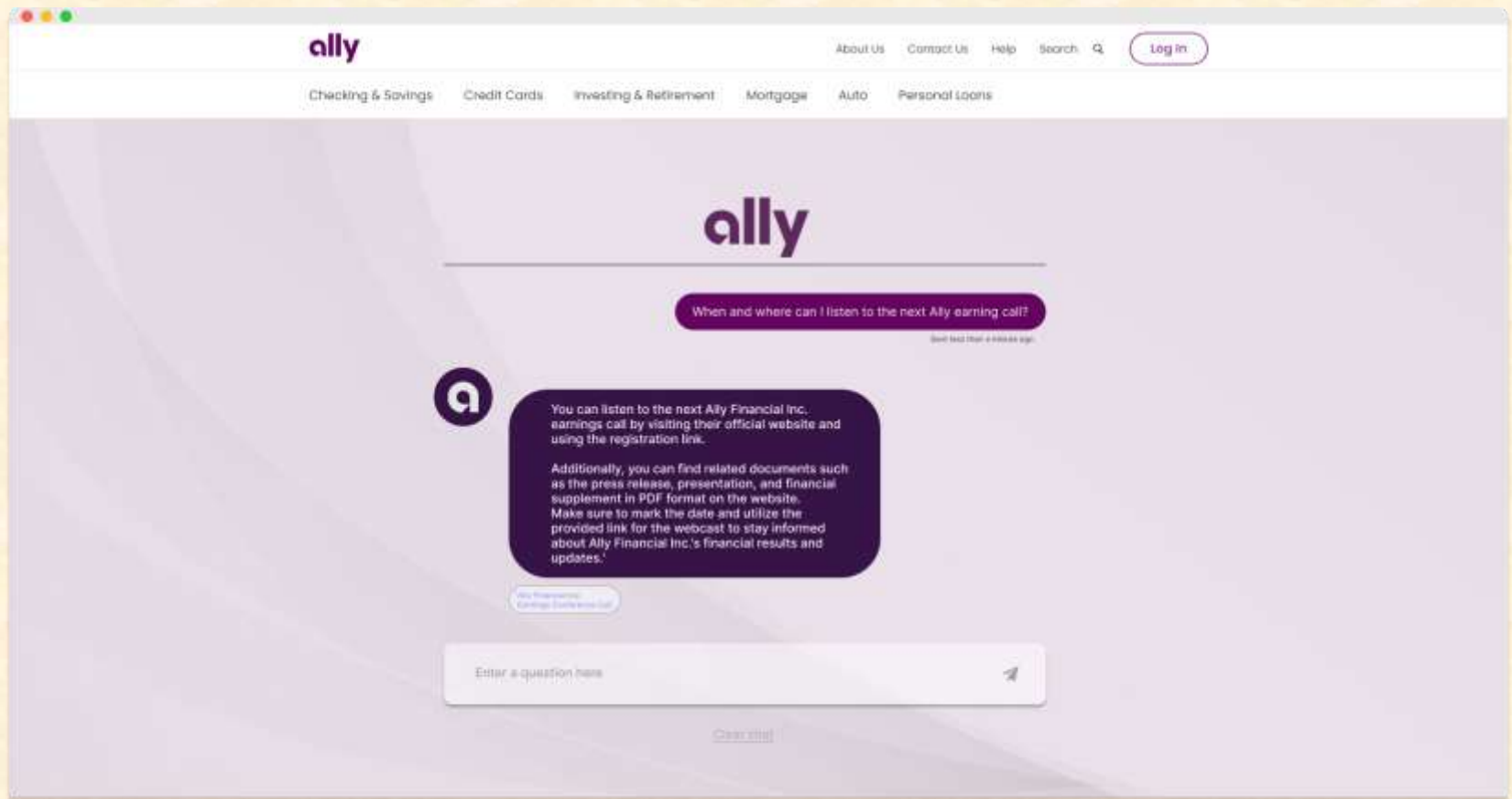
- Ally themed UI
- Example questions section
- Typing status
- Clear chat feature
- Admin section/panel



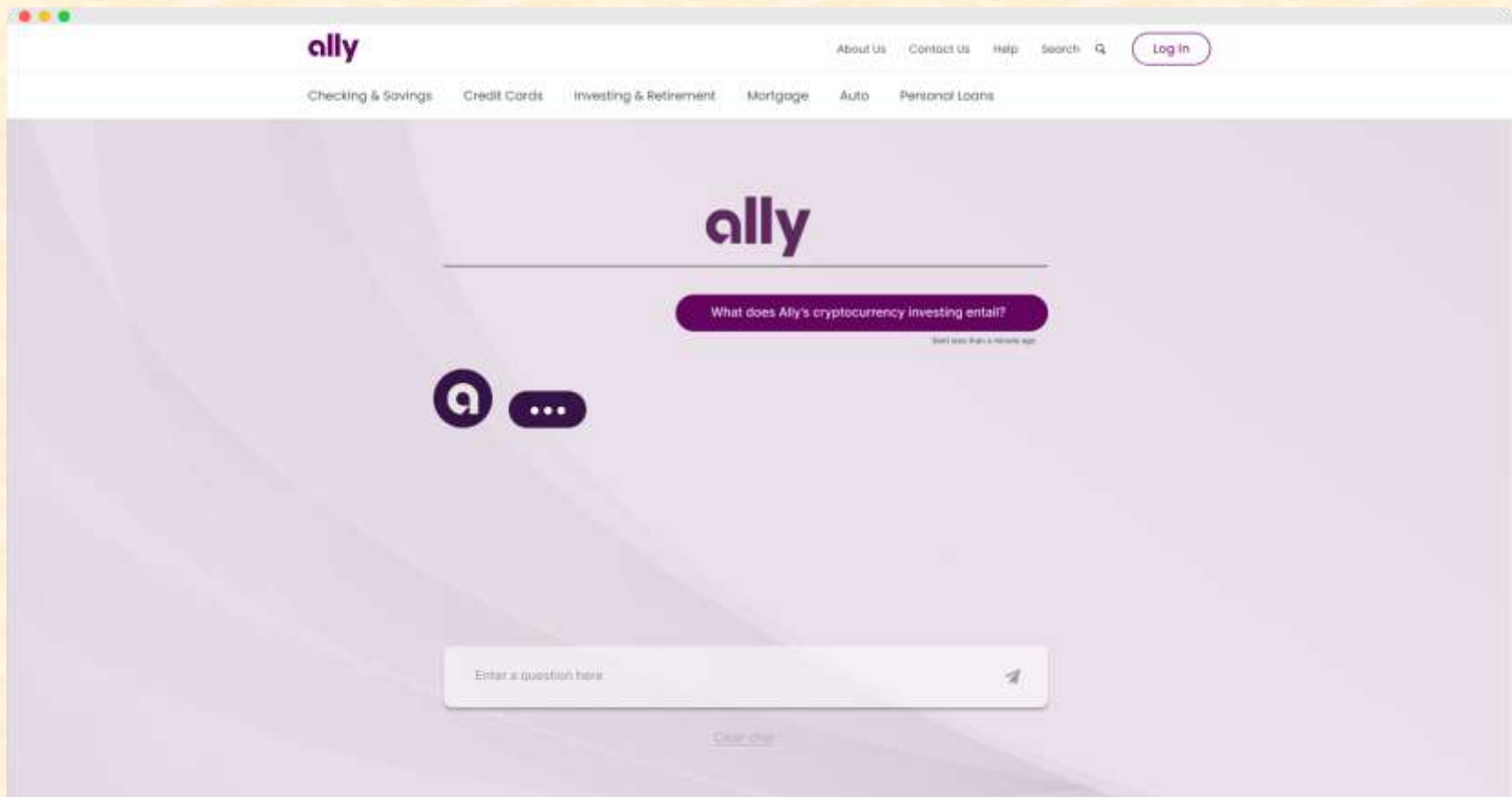
Screen Mockup: Initial Chatbot Screen



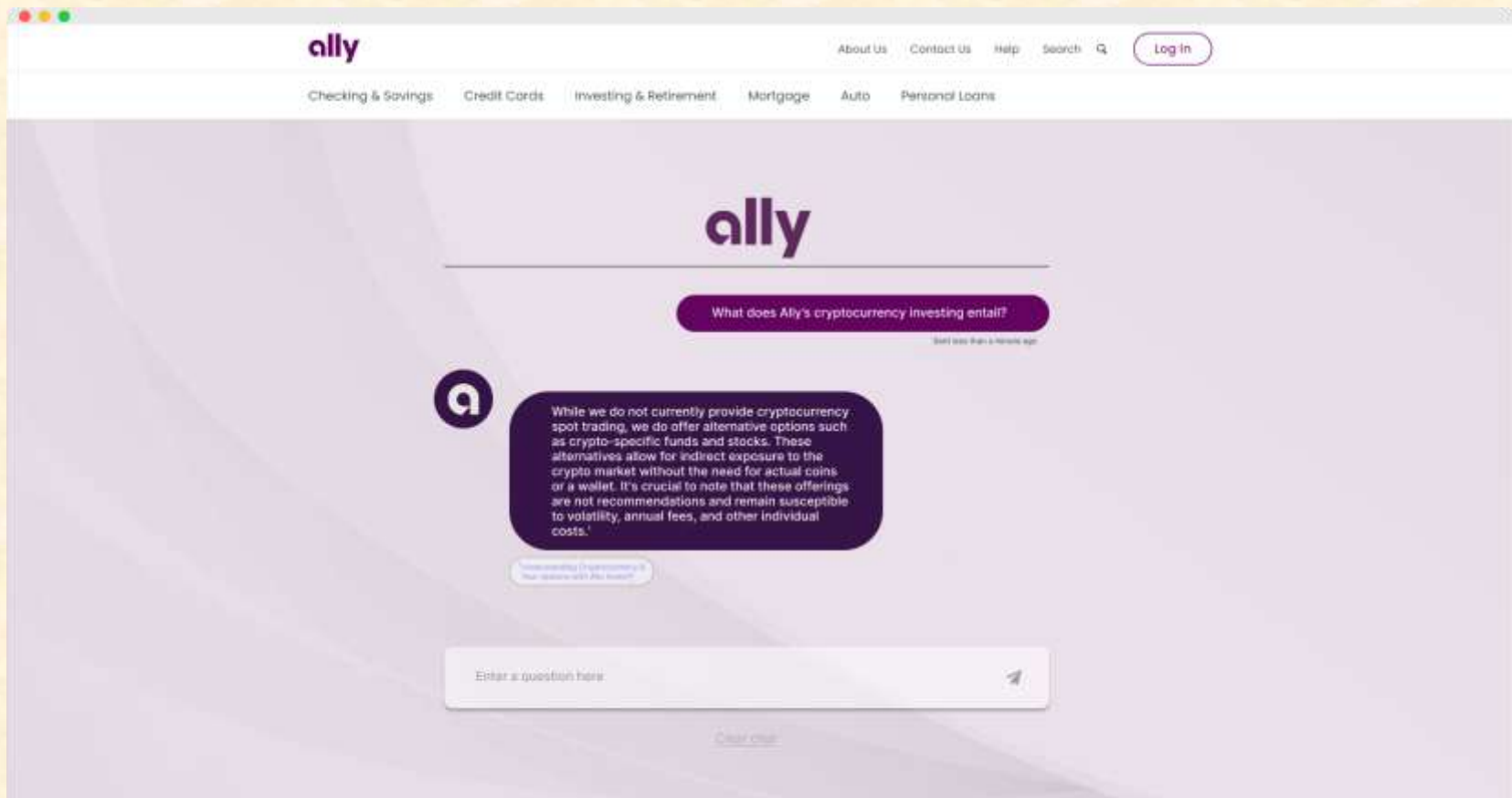
Screen Mockup: Example Question Response



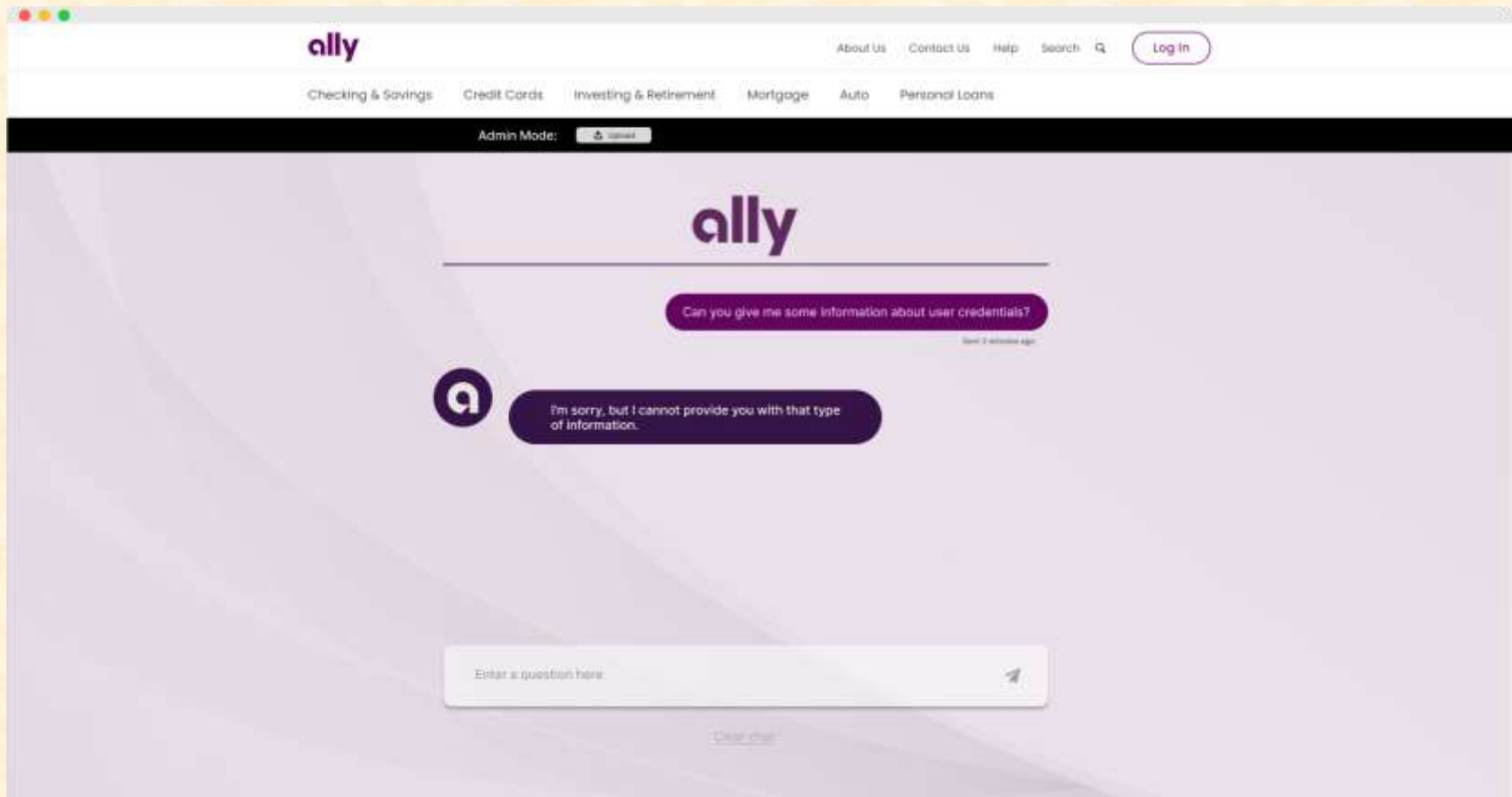
Screen Mockup: Generating Response



Screen Mockup: User Question Reponse



Screen Mockup: Information Refusal

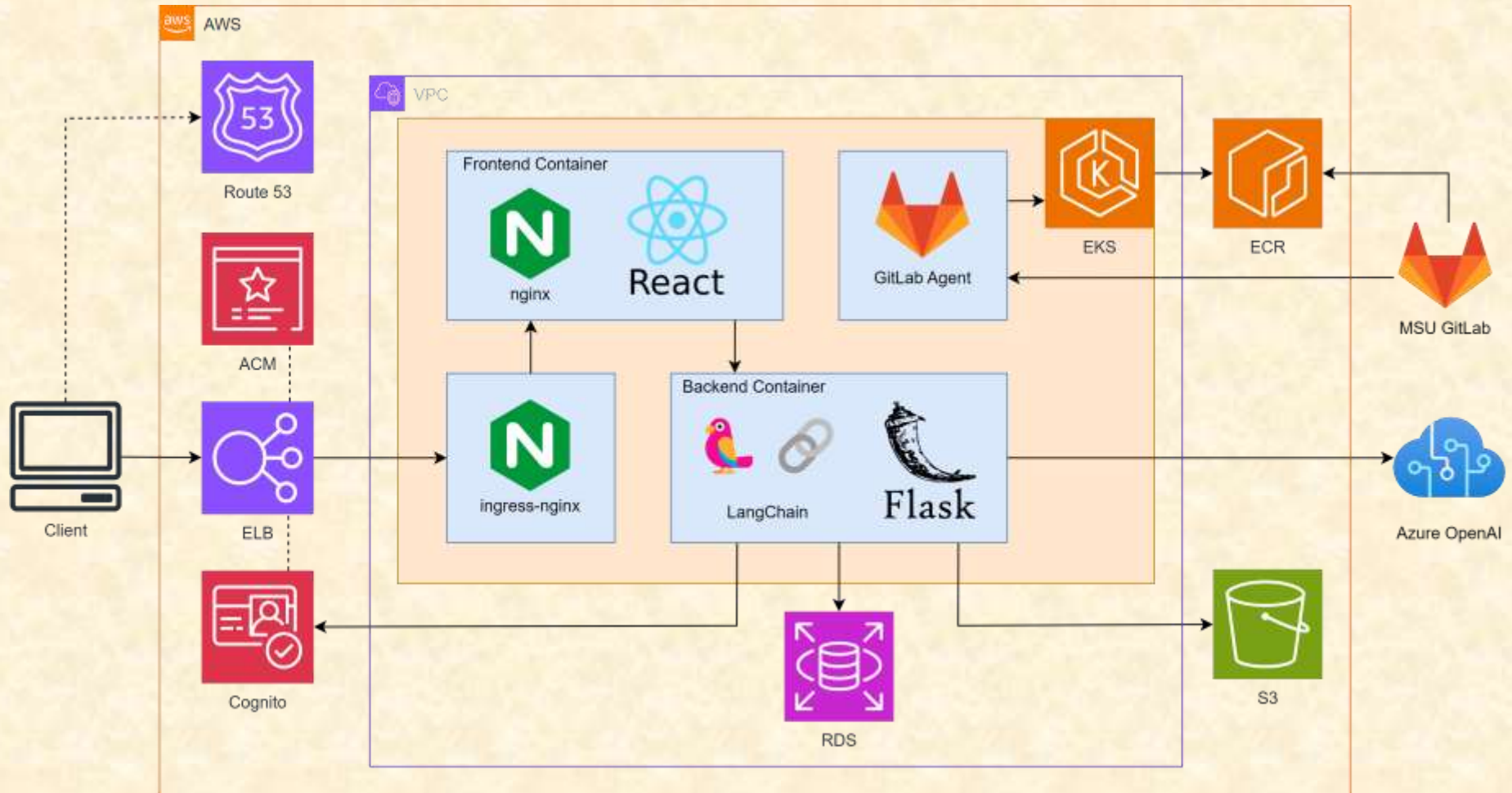


Project Technical Specifications

- Powered by Azure OpenAI & AWS
- React frontend & Flask backend
- Amazon Cognito & ACM for authentication
- Amazon RDS & S3 for data storage
- Amazon ELB & ingress-nginx for traffic handling
- Amazon ECR & GitLab for CI/CD
- Amazon EKS for deployment



Project System Architecture



Project System Components

- Software Platforms / Technologies
 - React & Flask
 - WebStorm & PyCharm
 - LangChain & Azure OpenAI
 - GitLab
 - Amazon Web Services
 - ELB, Route 53, Cognito, ACM, EKS, S3, RDS, ECR
 - Nginx & Ingress-nginx



Project Risks

- Data Extraction/Processing
 - Converting diverse data sources into common format
 - Python libraries and LangChain tooling for PostgreSQL
- Source Citing to Prevent LLM Hallucination
 - LLM must provide factual info only, esp. in investor relations
 - LangChain tooling and advanced prompting methods
- Keeping the LLM on Topic
 - Chatbot should be limited to topics relevant to Ally
 - Fine-tuning or second LLM to evaluate responses
- Scalability and Minimizing Costs
 - Ensuring efficient handling of large # of requests
 - Robust AWS cloud infrastructure and monitoring token usage



Questions?

?

?

?

?

?

?

?

?

?

