

**MICHIGAN STATE**  

---

**UNIVERSITY**

# Project Plan Presentation

## Audit Automation Tool

### The Capstone Experience

#### Team United Airlines Quality Assurance

Shafkat Kabir

Adam Collier

Emily Goldwater

Tejas Singhal

Haoyun Wu

Department of Computer Science and Engineering

Michigan State University

Fall 2023



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

# Project Sponsor Overview

- United Airlines has the most comprehensive route network and most fuel-efficient fleet among North American carriers
- They fly to hundreds of destinations across the world, with thousands of daily departures
- They employ over 70,000 people



# Project Functional Specifications

- Leverage AI to make auditing process automated for United Airlines
- By creating a web application to compare FAA and United Airlines regulations to vendor servicing manuals
- Showcase missing standards not found in the vendor manuals to streamline auditing process
  - reducing manpower required for auditing as well as the elongated auditing process time - bringing capacity back to the team

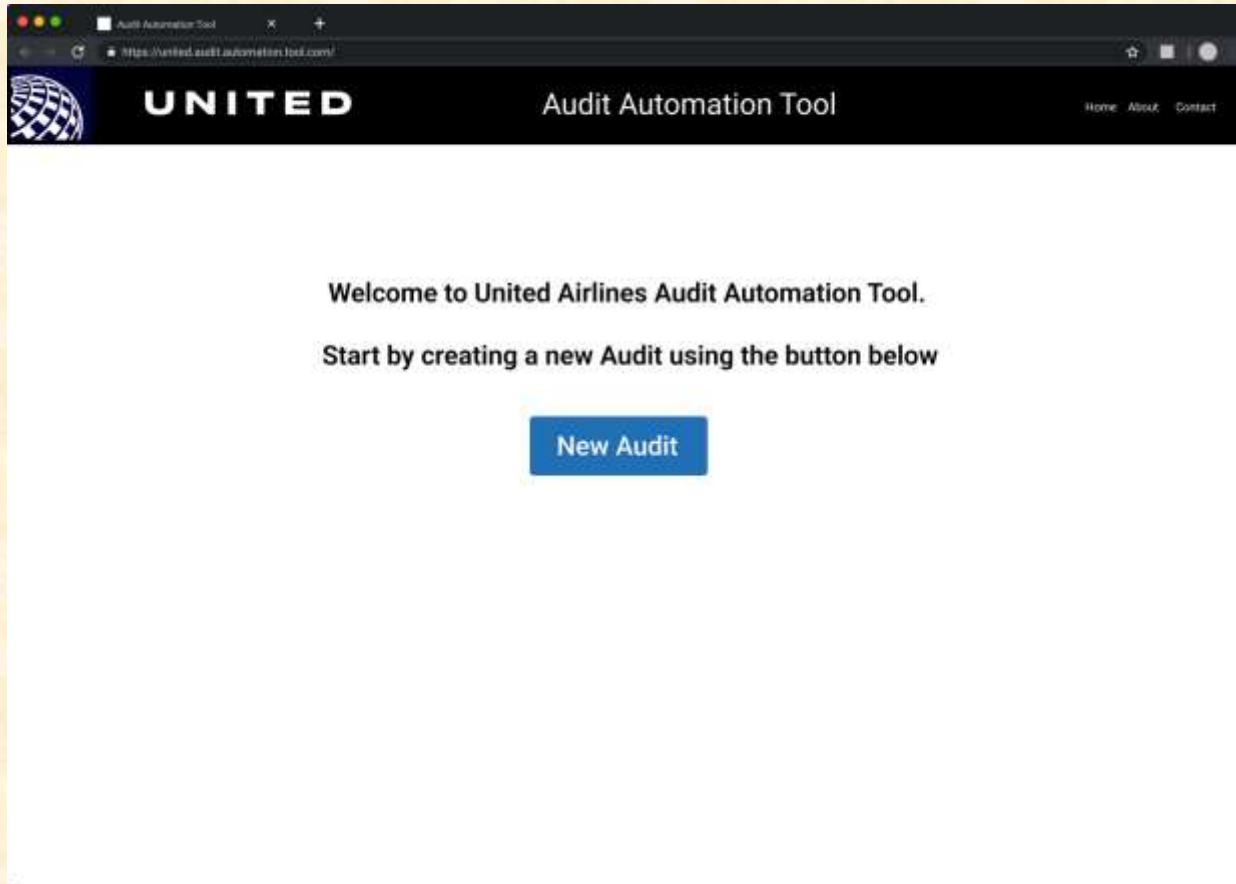


# Project Design Specifications

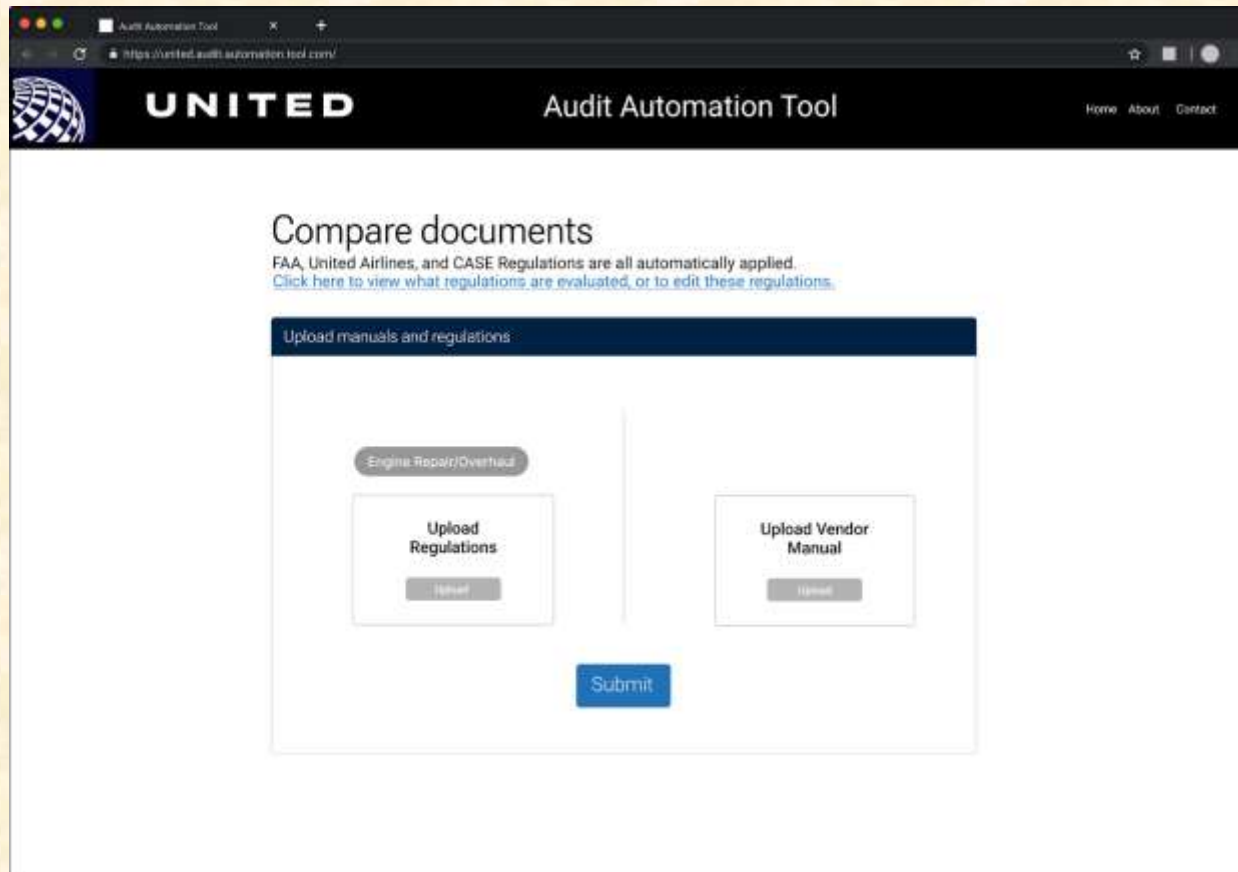
- App is a standalone web application
  - Compares supplied United Airline checklists and manuals and federal regulations to vendor documentation
  - Structure checklist and summary in an insightful way
- Follows United Airlines branding standards
- Four pages include: home page, document upload page, document selection page, and a results page



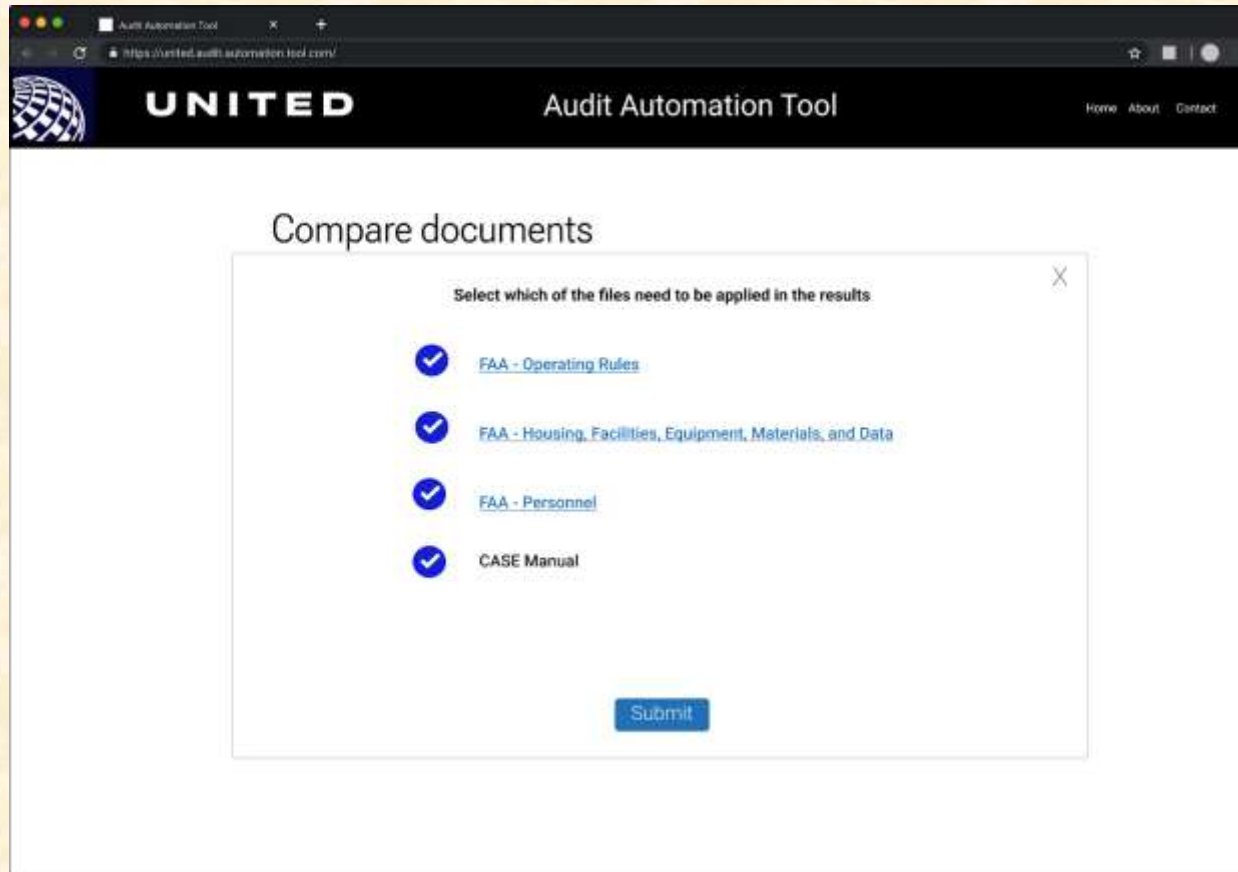
# Screen Mockup: Home Page



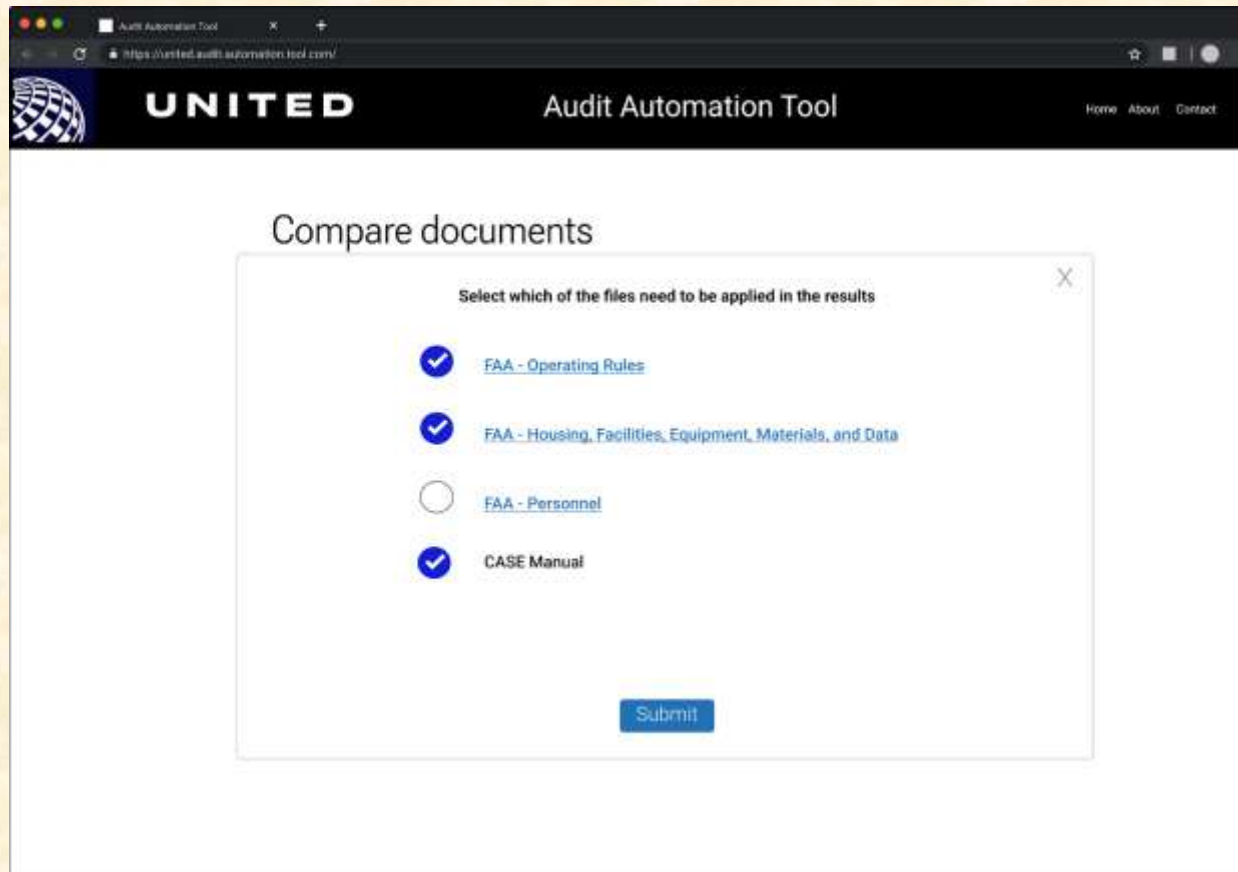
# Screen Mockup: Document Upload Page



# Screen Mockup: Document Selection Page

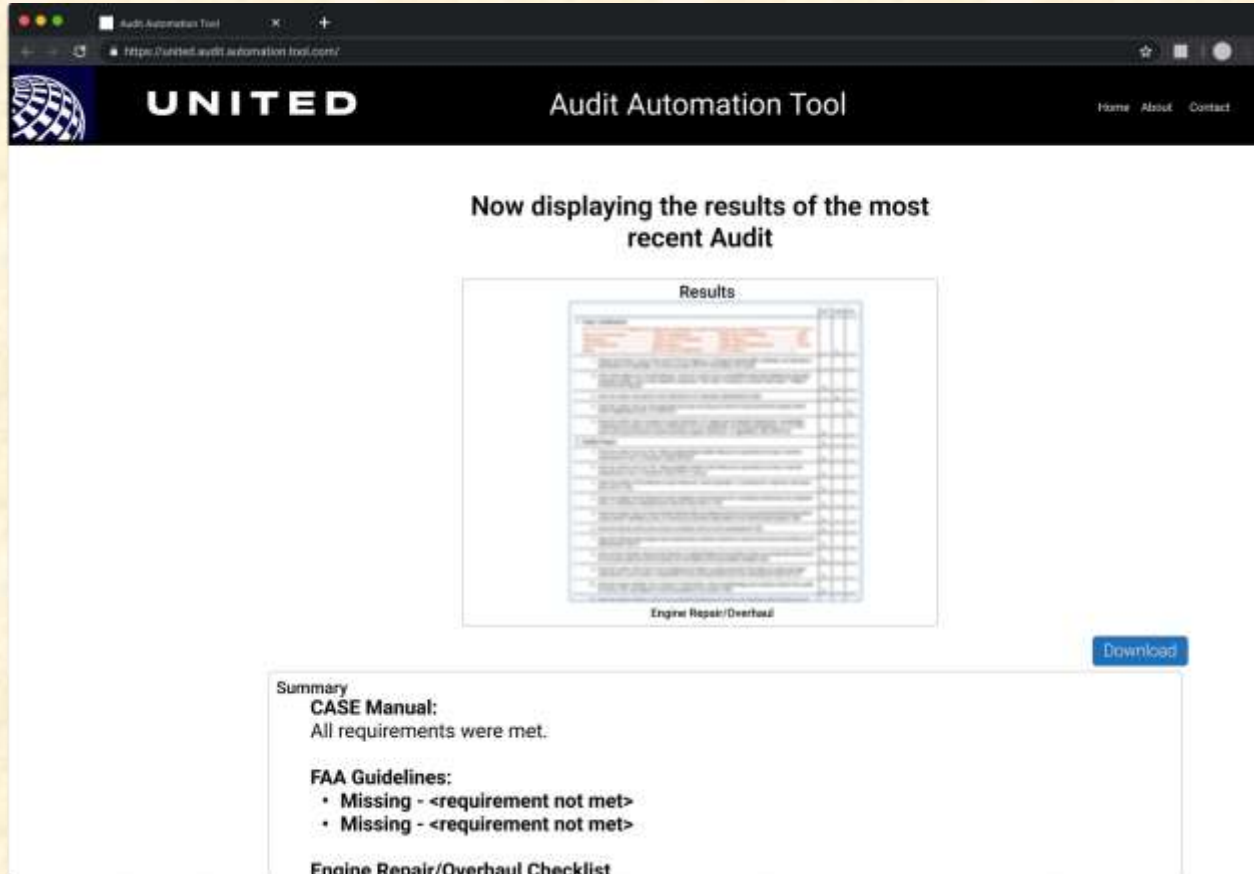


# Screen Mockup: Document Selection Page





# Screen Mockup: Results Page

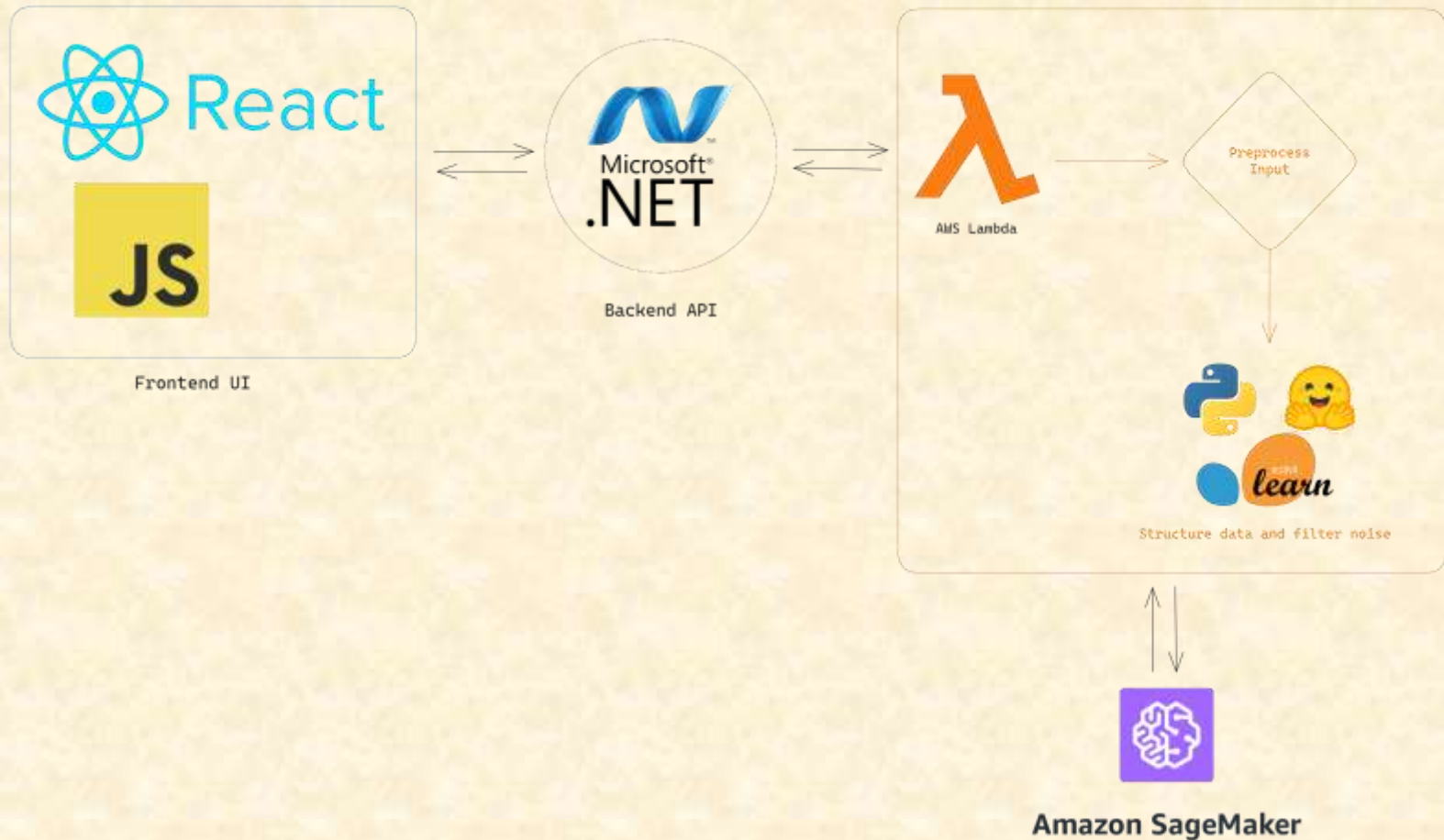


# Project Technical Specifications

- AWS
  - AWS ECS – Container service for website
  - AWS SageMaker – Machine Learning Service
  - AWS Lambda – Computing service for preprocessing
- C#/.NET – API for frontend-to-backend communication
- React.js – Fronted UI
- Python
  - Scikit-learn - ML library
  - Pandas - utilize to store and analyze data
  - Simplemma – lemmatize the word (i.e. jumping to jump)
  - HuggingFace Transformers – for generative AI model



# Project System Architecture



# Project System Components

---

- Software Platforms / Technologies
  - AWS
  - C#/NET
  - React.js
  - Python



# Project Risks

## Risks

- Risk 1

- Creating an efficient backend as the application uses big data
- Effectively preprocess the data, eliminating irrelevant data from the training set to decrease processing time and space
  - E.g., Addresses, names, titles, etc.

- Risk 2

- Understanding what noise to filter out from the data so that the important datapoints remain
- Communicate with sponsors to understand what portions of the documentation are necessary for evaluation

- Risk 3

- Creating a highly accurate model strictly using the data provided to us
- Create synthetic data using the provided data

- Risk 4

- Restructuring pdf/text manuals as a structured training set for the machine learning model
- Communicate with sponsors to understand the auditing process in order to learn how to preserve and properly structure the required data for the model



# Questions?

---

?

?

?

?

?

?

?

?

?

