MICHIGAN STATE
U N I V E R S I T Y

Project Plan Presentation
Intelligence Parser and Knowledge Assistant

The Capstone Experience

Team Roosevelt Innovations Knowledge Science

Jada Clark

Riley Cook

Brenden Cwiek

Maanasi Ganapathy

Austin Sawicki

Kyle Raeside

Department of Computer Science and Engineering

Michigan State University

Fall 2024

From Students…
…to Professionals

Project Sponsor Overview

- Subsidiary of Delta Dental of Michigan
- Software solutions company
- Creates insurance processing software

Project Functional Specifications

- Time consuming document transfer between insurance companies
- Benefit certificate translation
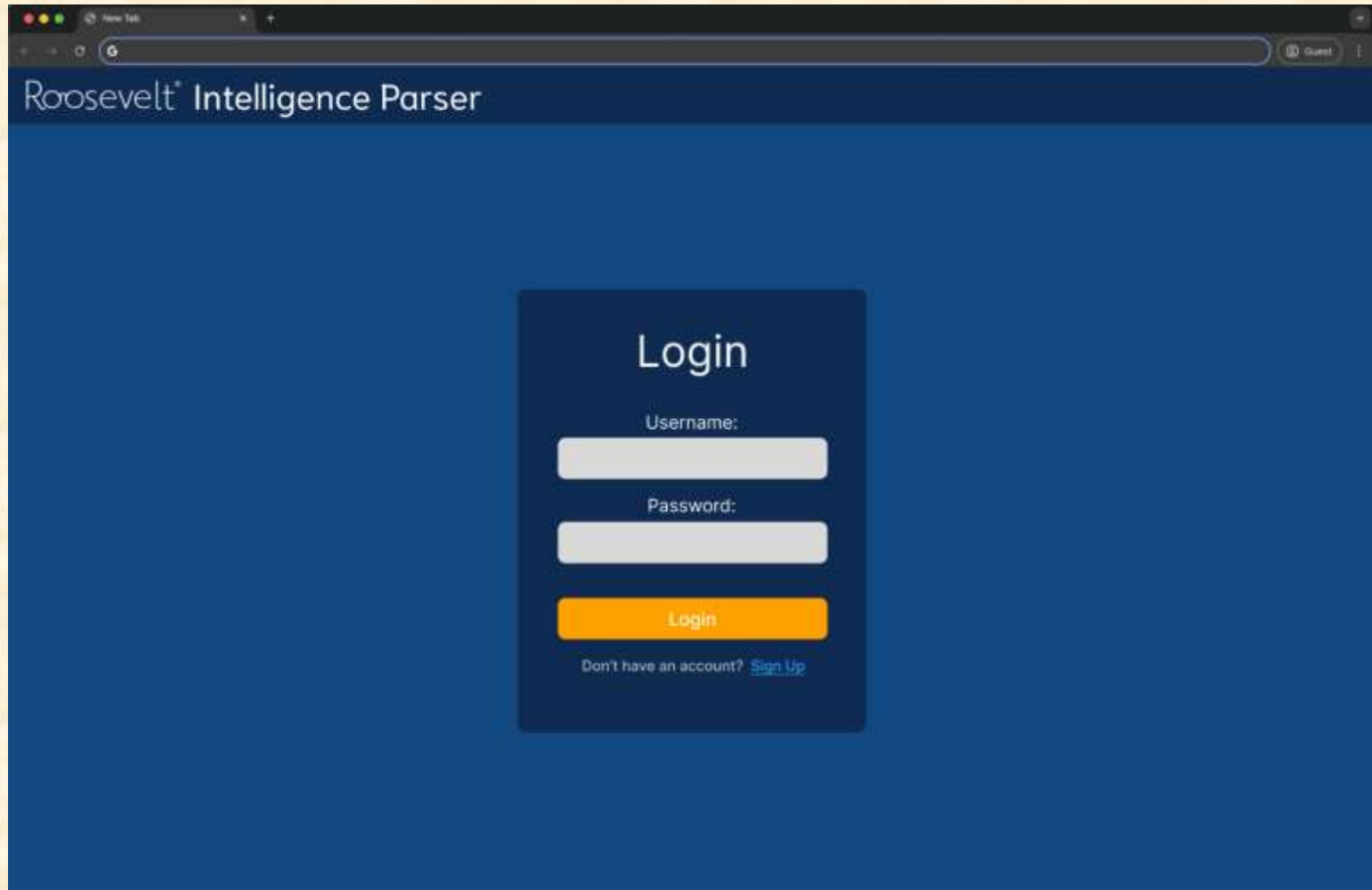- Find important benefit information
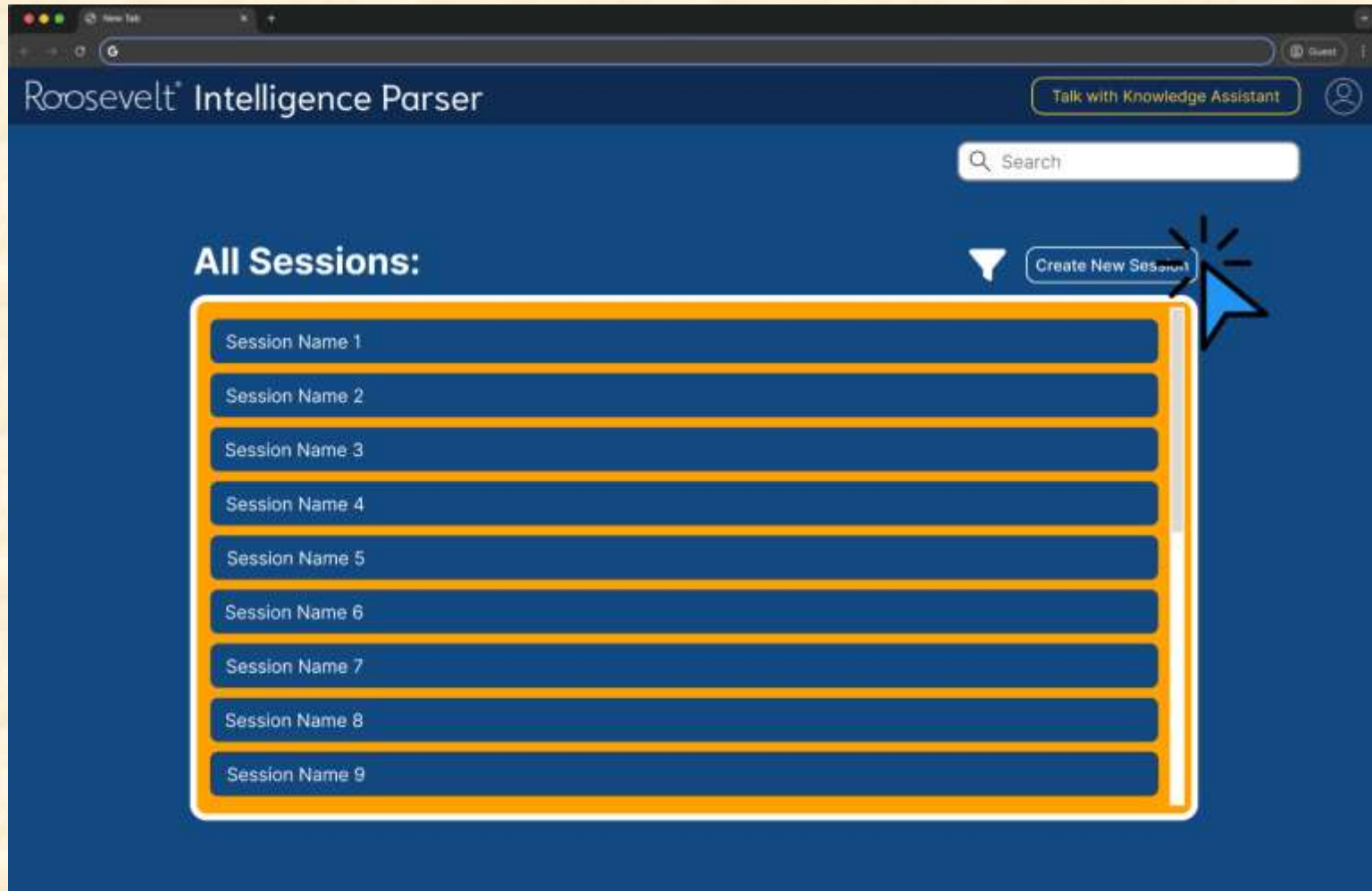- AI Chatbot

Project Design Specifications

- Dashboard Page
  - List of all uploaded Benefit Certificate PDFs (Sessions)
  - Upload a new PDF and create new Session
- Parsed Data/Validation Page
  - Have a page to display all parsed information
- Knowledge Assistant Window
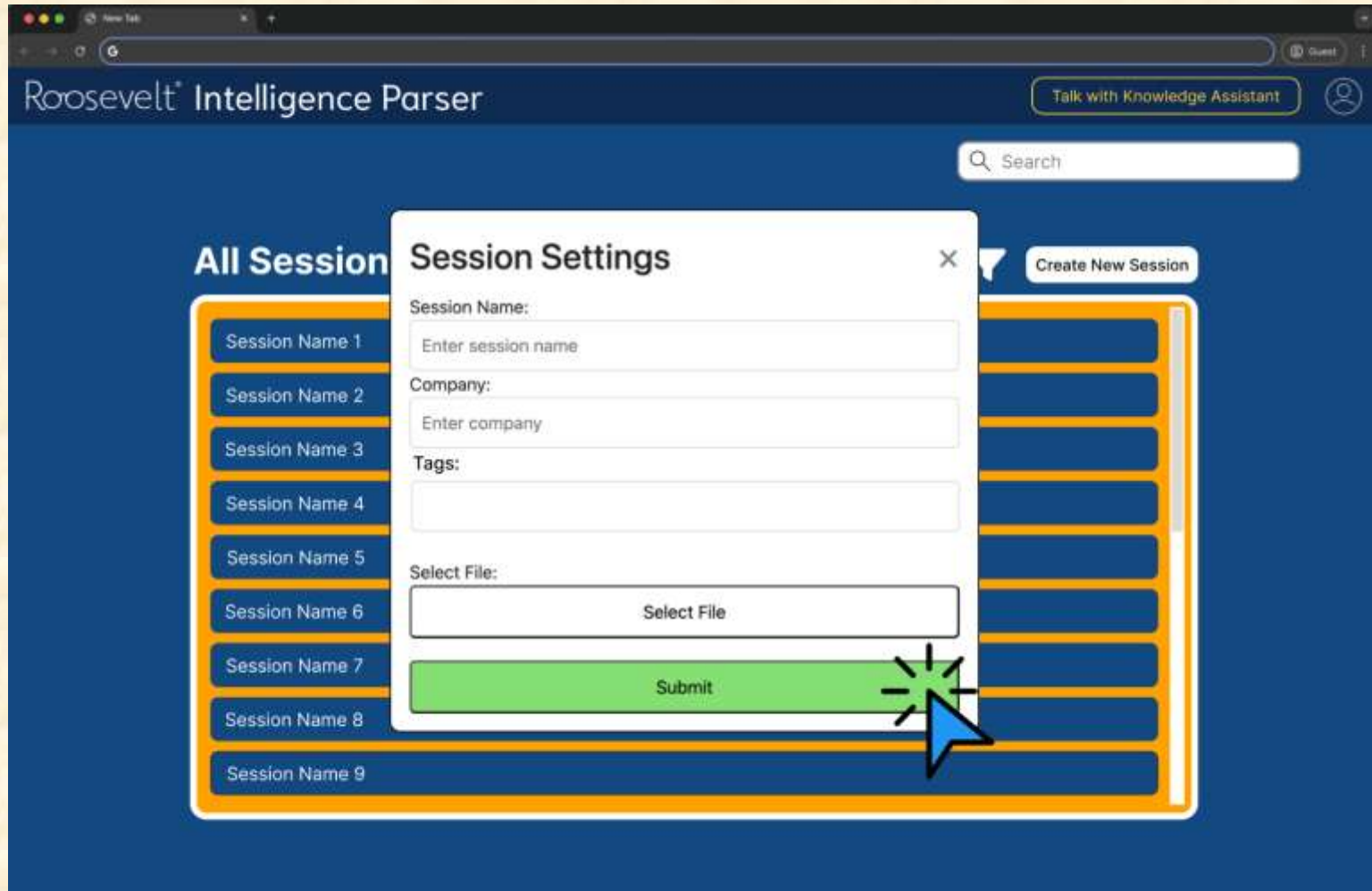 - Chatbot that answers any questions
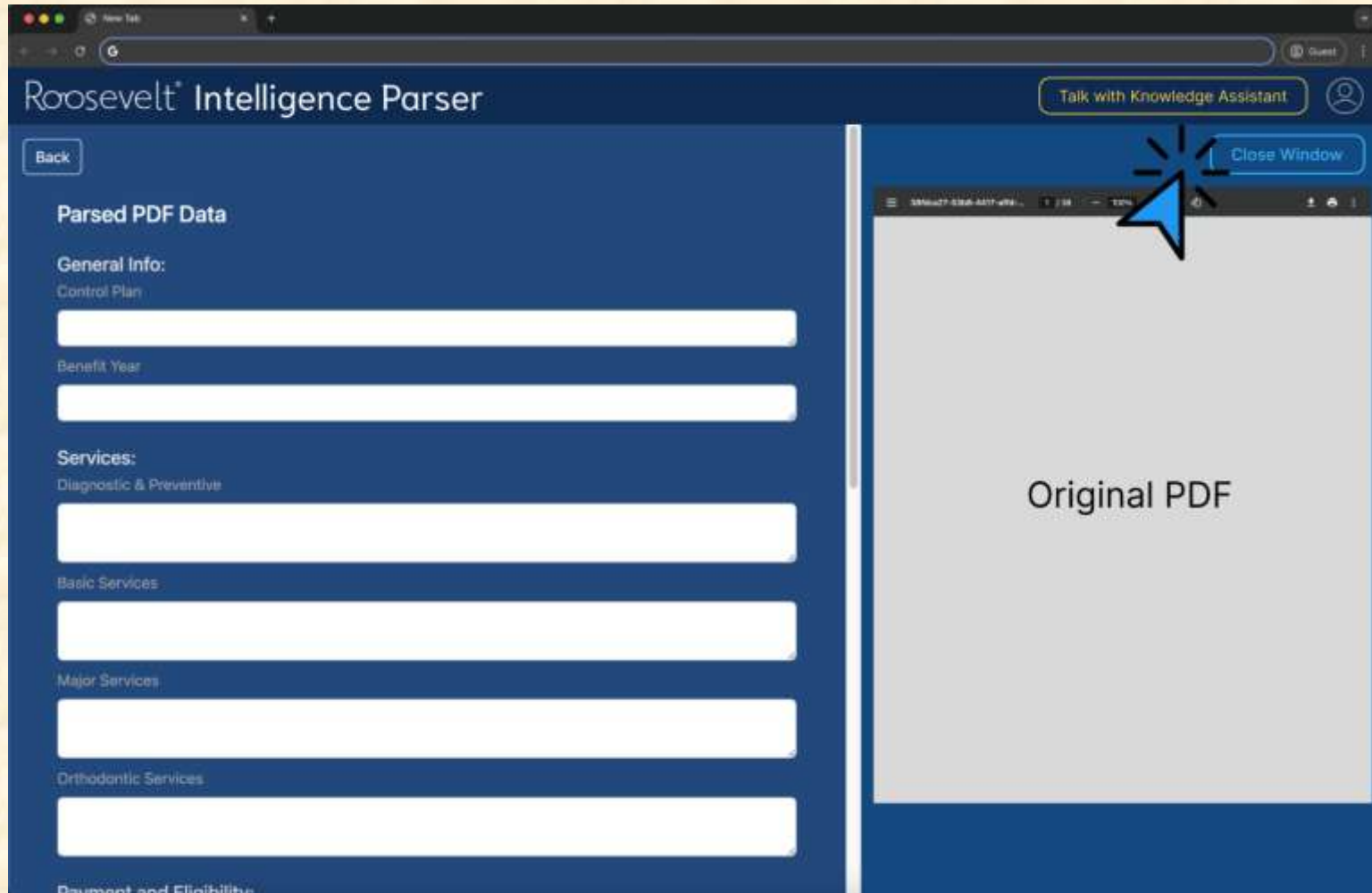
Screen Mockup: Login Page

Screen Mockup: Dashboard Page

Screen Mockup: Session Modal

Screen Mockup: Validation Page

Screen Mockup: Validation Page
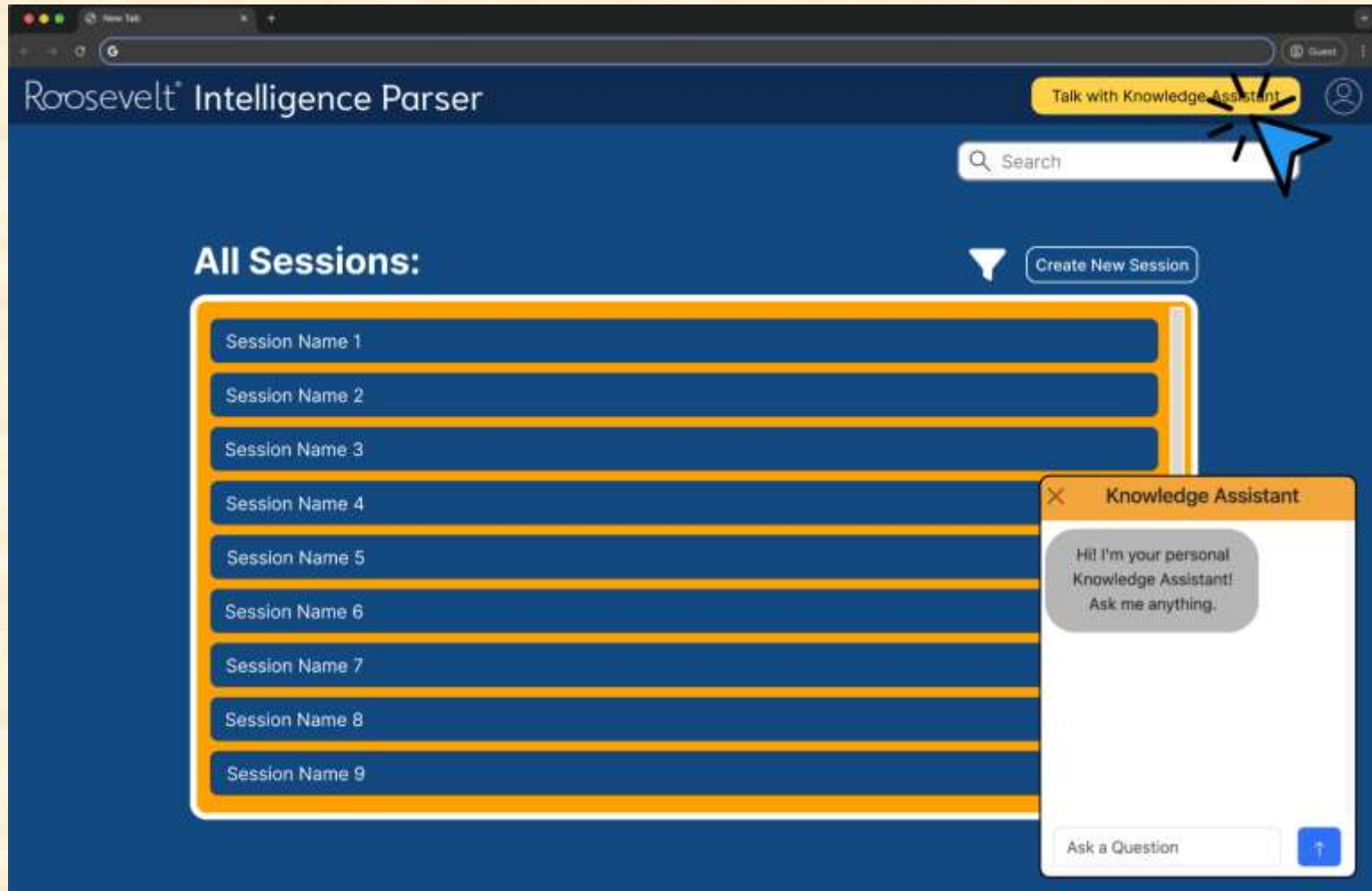
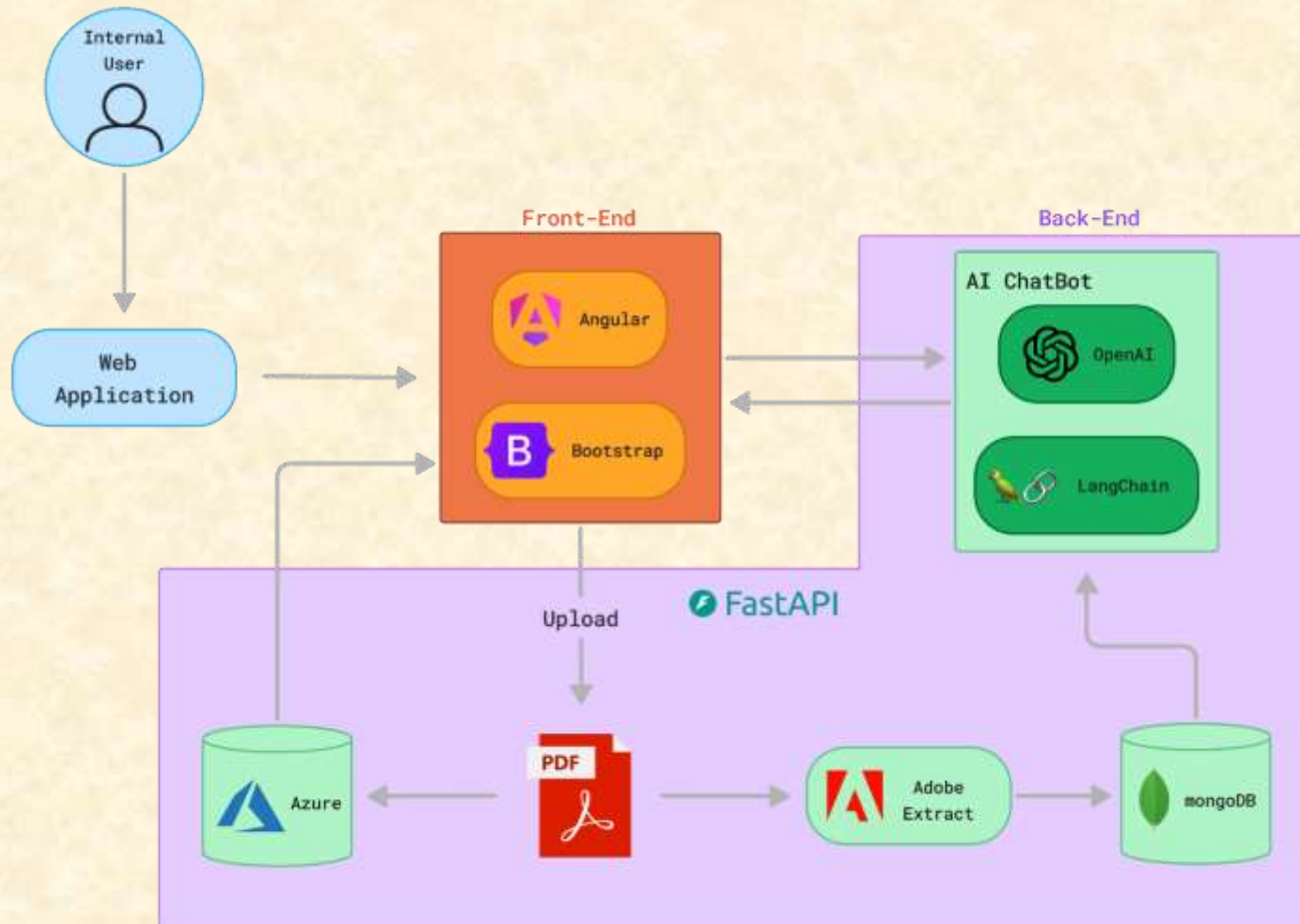# Screen Mockup: Filtering
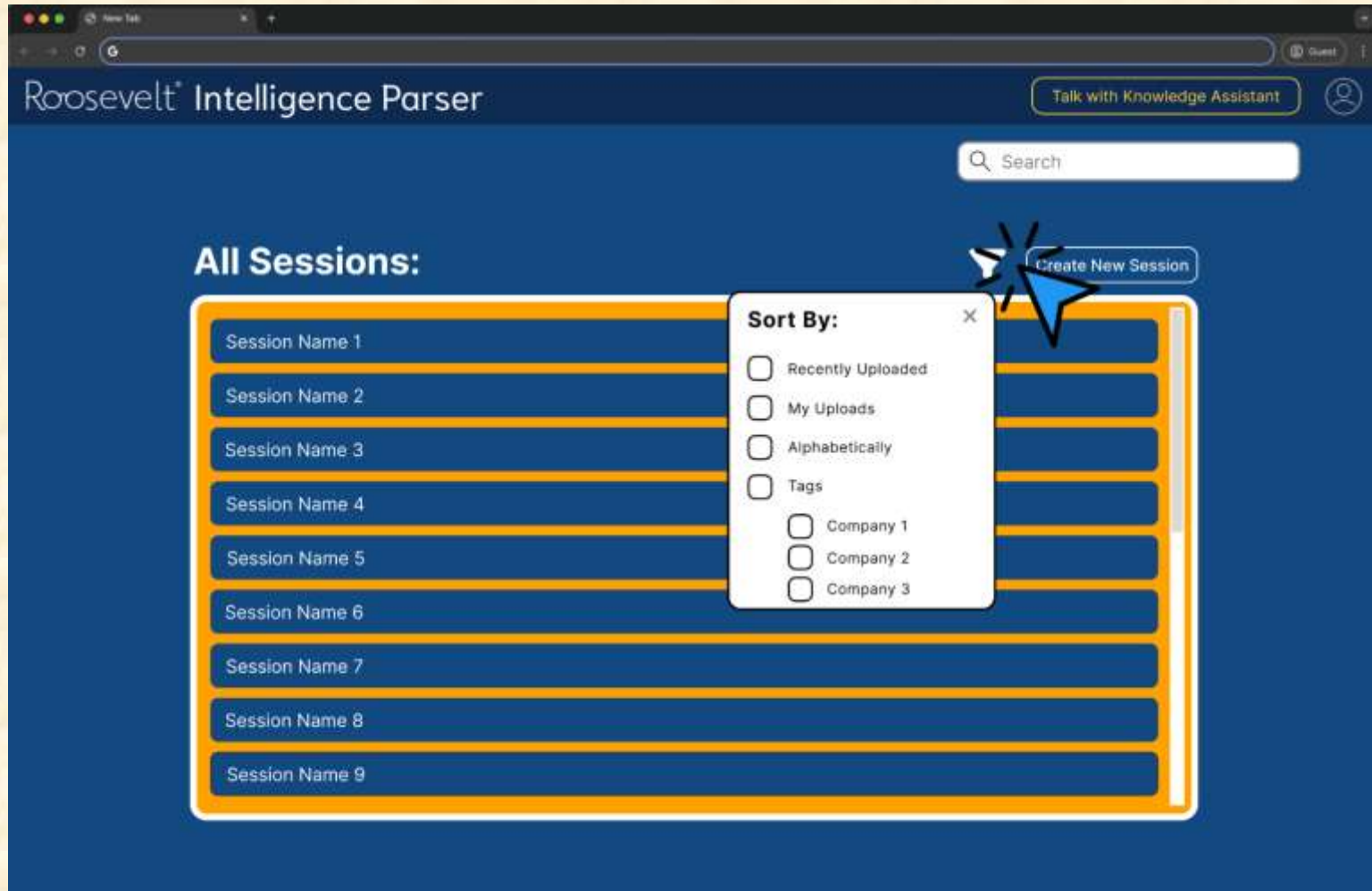
Screen Mockup: Chatbot

Project Technical Specifications

- Front-end
  - Angular (TypeScript)
  - Bootstrap (HTML, CSS, JS)
- Back-end
  - FastAPI framework
  - Azure BLOB Storage container for raw PDF data
  - MongoDB database for important parsed PDF data
  - Adobe Extract
- AI Chatbot
  - OpenAI
  - LangChain

Project System Architecture

Project System Components

- Hardware Platforms
  - N/A
- Software Platforms / Technologies
  - Angular Framework
  - Fast API Framework
  - MongoDB Atlas
  - Azure Blob Storage
  - Adobe Extract API
  - OpenAI API
  - LangChain Retrieval-Augmented Generation API

Project Risks

- PDF Differing Formats
  - PDF format not consistent between companies
  - Utilize natural language model to isolate correct data
- Generating PDF Chunks for LLM
  - PDF metadata needs to be properly and efficiently stored in database
  - Organize non-textual elements. Attach as metadata to parsed texts
- Efficient Database Schema
  - Difficult storing PDF session embeddings with user chat sessions
  - Create schema attaching chunks of data to user sessions

Questions?