✨

AI

At Focused Labs, we are working on building a custom domain-specific AI chatbot.

Scope:

Customized AI ChatBot that accelerates product development by making disparate and complex information easy to find; unblock teams to focus on what matters - building working software - rather than chasing down people and documentation.

Deliverables

- AI ChatBot tool and demo to demonstrate technical AI capabilities

- Repeatable, scalable process for building proprietary AI ChatBot for clients

▼ 🧪 Hypothesis

At large companies, teams find it challenging to locate the resources they need to build products quickly and effectively. Navigating complex legacy infrastructure and enterprise documentation is time-consuming, confusing , and just plain hard. Knowledge often resides with a single subject matter expert.

Building software does not have to be complicated. An AI Enterprise Knowledge Hub will speed up development time by bridging communication gaps and breaking down knowledge silos.

AI Enterprise Knowledge Hub will aggregate internal company information, paired with data available publicly, to provide faster access to information and thus faster development times.

Possible Outcomes

- Reduce the presence of the "lone SME" or "bus count of 1"

- Reduce duplicate efforts

- Reduce the time it takes to bring new products to market

- Build and deliver more modern products and experiences

- Reduce onboarding time for new developers

Personas

- **Target User:** development teams

- **Target Customer:** IT, Business or Technical Leader responsible for modernizing

▼ Goals and Milestones

- **Goals**

  - Establish Focused Labs as a knowledgeable AI partner in the broader AI tech space

  - Develop expertise in the generative AI content space

- **ChatBot**

  - Build an AI chatbot

    - For Focused Labs employees. Sample Q&A: As a user, I want to ask questions and get real answers. For example, what is our travel policy? what is an anchor? where can I find the town hall recording?

    - For Focused Labs clients. Sample Q&A:

  - Publish AI chatbot as a working demo on the Focused Labs website as evidence of technical AI capabilities

- **Publications**

  - Write and publish 3 to 5 blog posts that demonstrate best practices for building an AI chat bot and overall learnings about building in the AI space

- **Capability**

  - Draft process and approach for how Focused Labs builds Custom AI Enterprise Chatbots (or Knowledge Hubs) for customers

- Meet with at least 2 customers to demo AI capabilities. Customize pitch for a use case in their problem space.

Best Practices

Testing Best Practices

1. Create a large set of questions from our use cases (~50)

    - Establish must answer questions for our chatbot

    - Covers the breadth of data that we are ingesting

2. Automate asking questions and log answers

    - Asking our chatbot 50 questions takes ~5 minutes

3. Score Answers

    - Based on what the answers should be, not on the data fed into the model

    - Include a penalty score for when the app produces an error as a response

Cleaning Best Practices

- We wanted to automate and re-evaluate our data cleaning techniques as best practices to both improve accuracy and efficiency.
    - Lemmatization: algorithmic process of determining the lemma of a word based on its intended meaning, ex: ran, runs = run
        - Used python's Natural Language Toolkit library to implement
    - Remove unnecessary characters like emojis, html tags, etc…
    - Remove contractions
    - Remove punctuation

Development Best Practices

- Do not train or fine-tune the AI. Instead use vector embeddings.

  - Embedding, you are adding context to your prompt. This means that you'll need to store connections in a database, but will keep your information out of a model.

  - Cheaper and more likely to return specific pieces of information

- For data ingestion, leverage Llama Index. Their ecosystem, Llama Hub, has many open source data loaders that you can leverage.

- For querying data, leverage LangChain. Using the power of agents with the vector database as a tool, you'll be able to handle a robust set of questions.