Speak with your data – AI-QnA-App2

App building project using AI

1. **About me**
2. I am a SF-based business leader with 20+ years in enterprise tech GTM, with a special focus on strategic value selling.
3. I am a semi-technical person who knows basics on entity relationships / databases, python, workflows, APIs, etc.
4. I envision myself to be a builder-leader who can build tools / apps for personal and professional use using AI.
5. **Purpose of this project**
6. This app will help the user have a natural language chat with any database the user chooses to connect the app with. To start with we can make it work for a CRM-type database that contains information on accounts, opportunities, opportunity sizes, products, consultants, engagements, etc.
7. The app will also have other features such as news about a company which is an account (e.g. 10K, latest news)
8. **Efficient use of context window for development**
9. Please keep your responses short and to the point
10. Please refer to numbered items this document as necessary in your communication with me. If this document needs to be updated, please update it from time to time
11. Give me instructions one bit / one step at a time. Do not overwhelm with long outputs unless it is code
12. When I tell you something, ask any clarifying questions and check this document if required.
13. **Machines / folders on which I will do the development work**
14. Mac Mini M1 with an external SSD: “SSD”
15. MacBook Air: Macintosh HD – Users – venky
16. PC with NVIDIA GPU: folders need to be set up
17. **Use of Git and Google for working seamlessly across machines**
18. I would like to have all the coding files in Git so that I can work across these machines. My Git username is lvideation
19. I would like to store the CRM database in Google Drive but have it available “locally” on each machine while using the machine and then update it on Google Drive. When I move to another machine, I would like that machine to use the updated file but do it in a way that is local. Does this make sense?
20. **The database**
21. I have a version of this database in my folder under Macintosh HD / users / venky / AI-QnA-App2 / data / crmB.db
22. **Preferred folder structure**
23. I have a folder called AI-QnA-App2 created on my MacBook Air under Macintosh HD / users / venky.
24. **Secrets**
25. I would like secrets stored in an .env file
26. **My preferred development tools**
27. VSCode
28. Streamlit
29. Python
30. SQLite
31. Ollama / codellama
32. OpenAI Assistant API
33. OpenAI API
34. Gemini
35. **Development approach**
36. I would like to keep it as modular as possible so that it is easy to debug
37. **Functional Requirements**
38. General & Feature Requirements
39. The app should use Ollama / local LLM for queries related to the CRM database.
40. The app should prioritize using free APIs to get information on company / account for 10Ks and news
41. The app should have session management
42. The should have thumbs up / down for each interaction, with an ability for the user to provide feedback to the app
43. The app should have an admin panel where the user can see models and tokens used by session; a place where the user can provide context and glossary about the topic
44. The user should have the ability to connect to a different database. Once a new database is connected, the user should have the ability to upload a new context & glossary prompt. While a given database is connected, only sessions related to that database should be displayed. All previous sessions with any other databases should not be displayed.
45. Workflow Requirements
46. User selects new session or an existing session
47. User asks a question or enters any natural language text
48. OpenAI Assistant API reviews the request and routes it accordingly
    1. Request not related to the topic: Politely limit requests to the topic only
    2. Request related to CRM database: route it to Ollama. App provides an option to improve the way question is asked and refine it for efficient and effective SQL querying. This is just an option the user can choose to use or not. Once the user makes the choice and submits. Ollama responds with the answer; takes follow up questions and answers those follow up questions. If a follow up question requires external research, see C below.
    3. Request related to external ask (10K or news): use free API to get top 3 findings on business priorities from 10K and news; if there are follow up questions related to these, use OpenAI APIs but keep the response brief and focused on the topic
    4. Once the answer is provided, the message gets updated and updated in the thread / session
    5. User can rename a session if required
    6. User can delete a session
49. Feature Requirements
50. Admin panel
51. Select models
52. View token use by session
53. Create or update app context and glossary prompt