Interview Case: Gen AI Social Media Post Creator (PoC)

Objective

Develop a functional proof-of-concept (PoC) **web app** for creating LinkedIn-style social media image posts using **generative AI**. The solution must use **Streamlit, Gradio, or Dash** for the UI, and Python for backend logic. All AI components should use **open-source models** (e.g., Llama2, Mistral, BLIP2, Stable Diffusion, etc.). The app should support **multimodal RAG** (retrieval-augmented generation from PDFs, images, and web content) and include a **web search agent** powered by MCP or an agentic AI framework (e.g., Crew AI, AutoGen, LangGraph Agents, Langchain etc.).

Note: Models/Frameworks mentioned are just for reference and any open source model or framework can be used.

Core Features

1. Image Upload & Creation

- Users can upload their own images or generate images via an AI prompt (using open-source models like Stable Diffusion, with proper code and placeholders if no access).
- Optionally, users can generate images based on content extracted from PDFs or web search results.

2. Text Boxes with AI Assistance

- Each text box can be filled by:
    - AI-generated text (using a prompt, PDF content, or web search agent).

3. Multimodal RAG Integration

- Users can upload a PDF (e.g., resume, brochure) containing tables, charts,images, and text.
- System extracts relevant content from the PDF using RAG (LangChain, LlamaIndex, Haystack, etc.).
- Users can provide a web URL or search query; the system uses a **web search agent** (MCP or agentic AI framework) to retrieve and summarize relevant content.
- All retrieved content (text, images) can be used for post/caption generation.

4. Slideshow/Post Series

- Support for multiple slides/images in a post.
- Each slide/image allows for AI-generated or user-generated captions and descriptions.
- Allow rearranging slides and previewing the post as a sequence.

5. Preview & Download

- Preview the post as it would appear on LinkedIn.
- Option to download the final composed post (single image or slideshow as images).

Technical Stack

- **UI:** Streamlit, Gradio, or Dash (Python-based, rapid prototyping, supports multimodal inputs/outputs)
- **AI Layer:**
    - LangChain, LlamaIndex, Haystack, DSPy, or LangGraph for orchestration
    - LLM: Open-source (Llama2, Mistral, Falcon, GPT4All, etc.)
    - Multimodal models: BLIP2, LLaVA, Stable Diffusion, Hugging Face Vision Transformers
- **Image Generation:** Stable Diffusion or similar open-source models (API call or stub/mock if no access)
- **PDF Processing:** PyPDF, pdfplumber, or similar, integrated with LangChain for RAG
- **Web Search Agent:** MCP (Model Context Protocol) or agentic AI framework (Crew AI, AutoGen, LangChain Agents) for web content retrieval and summarization
- **Vector Database:** FAISS, Qdrant, Pinecone, or Weaviate (for embeddings and retrieval)
- **NLP:** Hugging Face Transformers, spaCy for text processing and entity extraction

Expected Deliverables

1. **Working Code**
    - Streamlit, Gradio, or Dash app (Python)
    - All core features implemented (functional with real AI models or proper placeholders)
    - Well-structured and commented code
2. **Documentation**
    - Clear README with setup instructions
    - How to insert API keys if available
    - Brief note on how you would productionize or scale this system
3. **Demo**
    - Short demo (screenshots, GIF, or short video) showing main features and workflow

Bonus (Optional but adds value in evaluation)

- Authentication for users
- Multi-language support
- Docker setup
- Error handling and user-friendly messages
- Advanced multimodal features (e.g., image captioning, OCR, text-to-image with context from PDF/web)
- Integration with Azure AI Studio, Azure Machine Learning, or Azure Databricks (mocked if no access)

Evaluation Criteria

- Completeness and functionality of the PoC (core features met)
- Code quality, organization, and documentation

- Clarity and usability of the UI/UX
- Proper AI integration or clear placeholders
- Approach to multimodal RAG (PDF, web, images) and AI content generation
- Any extra innovative or useful feature added not mentioned

Timeframe

You have **4 working days** to complete and submit your solution. Good luck! We are looking forward to seeing your creativity and technical skills in action.