Solution Architecture

Bite Finder

Process (People)

- Train/fine-tune vision-language models to classify bug bites
- Augment image data
- Synthesize user text data
- Build vector database of relevant medical documents
- Build frontend app
- User uploads image and types symptoms and location of bug bite
- User inputs are checked for quality
- User receives bug bite prediction
- User can select pre-made questions or ask their own to our bug expert

Execution (Code)

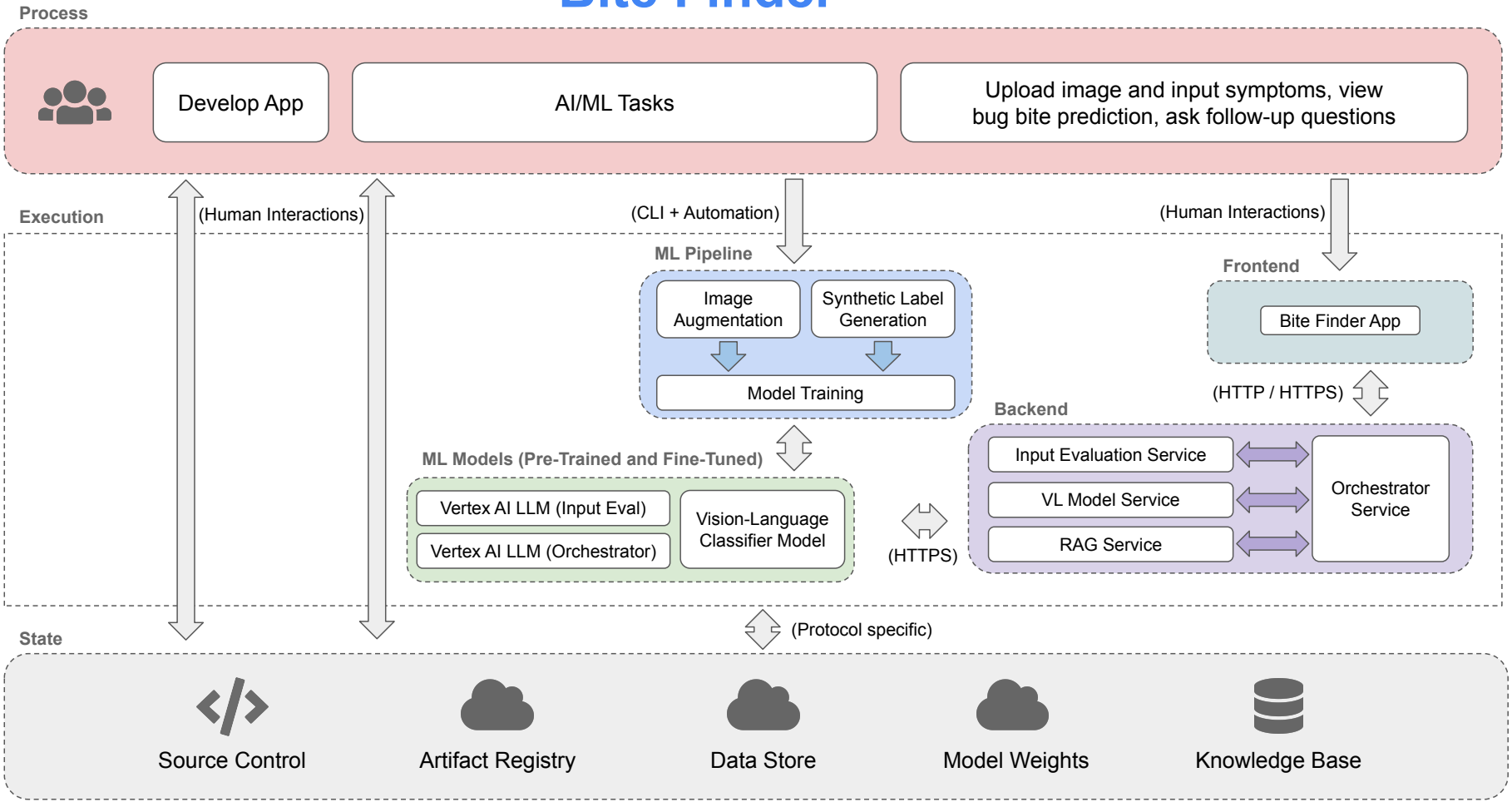
- Call LLM for input evaluation
- Feed image and text through processor and vision-language model
- Use the best VL model for prediction
- Agentic orchestration of LLM with RAG for follow-up Q/A
- Input evaluation, VL inference, and RAG set up as API services that orchestrator service calls

State (Source, Data, Models)

- Save dataset to common store
- Save Docker images to common store
- Save text chunks and embeddings of relevant medical documents
- Save model weights and metadata for loading at inference
- Experiment tracking
- Source code repo for collaboration
- Use pre-trained LLM for input evaluation and Q/A

Solution Architecture

Bite Finder



Technical Architecture