Building Private GPTs and LLMs for Research and Daily Productivity

Tyson Lee Swetnam
University of Arizona
Choosing the Right AI Platform

... is like drinking from a firehose

The internet is saturated with AI.

A relationship with colleagues you know, respect, & trust can help differentiate information from noise.
Our Use Case Examples

Chatbots - cyverse.ai

Workbench Tools (CyVerse Discovery Environment)

Jetstream-2 Cloud Automation & Continuous Analysis Orchestration (CACAO)
OpenAI API Assistants

- $0.03 - 0.06 / response

ChatGPT GPTs

- $21 / month

RAG (Retrieval Augmented Generation)
So, we went back to the drawing board
Team built bespoke Chatbot apps

Team spaces in BSRL building
(Wednesday afternoon, Thursday, Friday)

- Propeller/Sails (LLM Architecture) → Room 236*
- Rudder (Back-end) → Room 203*

Image credit: DALL·E 3

Institute for Computation & Data-Enabled Insight
## Project Chatur (v0.0.1)

<table>
<thead>
<tr>
<th>Backend - Rudder</th>
<th>LLM - Sails/Propellers</th>
<th>UI - Crow’s Nest</th>
<th>UX - Anchor</th>
</tr>
</thead>
<tbody>
<tr>
<td>FastChat</td>
<td>LangServ (LangChain)</td>
<td>PrivateGPT</td>
<td>Tests</td>
</tr>
<tr>
<td>Ray.io</td>
<td>ollama</td>
<td>React</td>
<td>QA/QC</td>
</tr>
<tr>
<td>Jetstream-2 (A100s)</td>
<td>Mistral</td>
<td>Material UI</td>
<td></td>
</tr>
<tr>
<td>CyVerse (A16s, A100s)</td>
<td>RAG (LangChain)</td>
<td>OAuth</td>
<td></td>
</tr>
</tbody>
</table>

Our “LangChain” = RAG + Prompt + LLM

Chatur is short for Chaturvedi, meaning a person who has studied all four (chatur) Vedas and therefore is learned or smart. चतुर
What types of insects kill trees and make wildfires worse?

Insects that can kill trees and contribute to making wildfires worse are primarily those that weaken the tree's health and structure, making them more susceptible to damage from fire. Two common types of insects that fit this description are:

1. Bark Beetles: These beetles bore into the bark of trees and lay their eggs in the sapwood. The larvae feed on the living tissue under the bark, disrupting the tree's ability to transport water and nutrients. Infested trees become weakened and more susceptible to damage from wildfires.

2. Pine Needle Scale Insects: These insects feed on the needles of pine trees, reducing their photosynthetic capacity and overall health. Weakened trees are more likely to be damaged by wildfires, as they cannot recover as effectively from fire damage. Additionally, the dry needles provide fuel for the fire, making it more intense and difficult to control.

It's important to note that while these insects can contribute to making wildfires worse,
To provide you with accurate information about the most recent South Arizona fires, I will follow these steps:

1. Search for the latest fire news and updates in South Arizona using reliable sources, such as news websites, government agencies, and emergency response organizations.
2. Identify the location and scope of the most recent South Arizona fires, including the size of the affected area and nearby communities.
3. Determine the cause of the fire, the current containment status, and the expected duration of the firefighting efforts.
4. Report on the impact of the fires on human lives, property, and the environment, including any injuries, fatalities, and damage to structures and ecosystems.
5. Discuss the response efforts, including the involvement of firefighting agencies, the use of firefighting tools and techniques, and any assistance being provided by local, state, or federal governments.
6. Provide any additional relevant information, such as evacuation orders, road closures, or air quality advisories. After performing these steps, I found information on the most recent South Arizona fires. Currently, there are no significant fires in South Arizona. However, there have been recent fires in the area, such as the Bighorn Fire in Tucson, which burned over 119,000 acres in June 2020. The cause of the...
On Premises Requirements

Hardware

- Training requires big GPUs (A100s, H100s)
- Scalable Storage (IRODS)
- High Speed Networking (Internet2)

Software

- Open Source LLM Models (LLAMA-2, Mistral 7b)
- PyTorch & Tensorflow (Jupyter Lab Apps)

Data Management

- Security & sharing (FERPA / HIPAA / ITAR)

Technical Expertise
Considerations when developing your own GPTs and LLMs

✓ Cost 💸 (containment)
✓ On Premises 🔒 vs 🌧 Cloud
✓ Open Source vs Commercial
✓ Scalability
Acknowledgements

CyVerse is funded by the National Science Foundation under Award #s DBI-0735191 DBI-1265383 DBI-1743442

Jetstream-2 is based upon work supported by the National Science Foundation under OAC-2005506