Information regarding the poster:  
*Poster Size: 36in x 24in, 300dpi*  
Introduction:  
- Customizable Ai-powered search agent  
- Tailored to financial usages  
Vision:  
- “Capturing” complexity of websites, understand the content and structure of websites, so the agent can explain and analyze them  
- Know the user’s habits: interests, hobbies, web surfing habits, goals, etc. via a dynamic database  
- Upgrade general public into professionals  
Agent Details:  
- Side-by-side dual-model output for user to compare results  
- Select models from Open Financial Leaderboard (OFLL) directly  
 - OFLL paired with FinGPT agent to serve the public——not just any agent, not just any leaderboard  
- Specify URL / API Endpoints for the agent to prioritize in accessing.   
 - Source-checking  
- Googling keyword extracted from prompts to enhance context  
 - Source-checking  
- Local file parsing via RAG  
- Run in air-gapped infrastructures within an institution  
UI / UX:  
- Currently existing in the form of a web extension  
- No fancy colors (black-white-grey to address color-blindness)  
- Draggable, closable, condensable

## Abstract

Current large language models (LLMs) have proven useful for analyzing financial data, but most existing models, such as BloombergGPT and FinGPT, lack customization for specific user needs. In this paper, we address this gap by developing FinGPT Search Agents tailored for two types of users: individuals and institutions. For individuals, we leverage Retrieval-Augmented Generation (RAG) to search local documents and user-specified data sources. For institutions, we employ dynamic vector databases and fine-tune models on proprietary data. There are several key issues to address, including **data privacy**, **the time-sensitive** nature of financial information, and the need for **fast responses**. Experiments show that FinGPT Search Agent outperform existing models in accuracy, relevance, and response time, making them promising for real-world financial applications. [Customized FinGPT Search Agents Using Foundation Models]

## Introduction

Large Language Models (LLMs) are revolutionizing the financial sector by enhancing data analysis and decision-making processes. However, existing financial LLMs (FinLLMs) like BloombergGPT and FinGPT lack the capability to provide customized advice to individuals and institutions, especially when handling proprietary or personal data. There is a growing demand for solutions that address unique user requirements while ensuring data privacy, real-time updates, and swift response times.

**Our Solution**

**We propose two customized FinGPT agents: one tailored for individuals and another for institutions.** These agents leverage Retrieval-Augmented Generation (RAG) and model fine-tuning to integrate data from diverse sources. This approach offers personalized financial insights while ensuring data privacy and timely responses. We also developed tailored Graphical User Interfaces (GUIs) to enhance user interaction for both user types.

**Key Challenges**

* **Privacy:** Securing sensitive financial information for both individuals and institutions.
* **Time-Sensitivity:** Efficiently processing and integrating the most up-to-date financial data.
* **Response Time:** Delivering high-quality responses within acceptable time frames (under a dozen seconds).

**Contributions**

* **Customized Agents:** Developed FinGPT agents that address unique user needs. It may run in an air-gapped infrastructure.
* **Tailored GUIs:** Designed and implemented dual-model interfaces for model comparison.
* **Enhanced Performance:** Demonstrated through experiments that our agents outperform existing LLMs in accuracy, data freshness, and response speed, highlighting their potential for real-world applications.