**AI-Powered HR Knowledge Agent: Streamlined Document Insights and Query Assistance**

**Description:** Build an agent that assists HR users by answering queries related to company Google Drive documents (Employee Handbook, Leave and Attendance, Friend and Family Policy & etc) and provides contextual information from company policy documents.

* **Key Features:**
  + **Fine-Tune the LLM:** Fine-tune LLM on HR policy summaries and frequently asked questions for general knowledge.
  + **Use RAG for Specificity:** Pair it with a document retrieval mechanism to handle granular or dynamically changing queries.
  + **Intelligent Query Answering:** Develop an agent to enable HRs to ask questions about company policies quicker for their operational needs and to remember HRs conversation
* **Steps:**
  + Collect a dataset of OneDrive documents.
  + Preprocess and clean the data to train the LLM on company-specific terminology.
  + Fine-tune an LLM to answer queries related to company policy related documents.
  + Implement RAGs to retrieve relevant data from the Google Drive in response to HR queries.
  + Build an agent interface for HRs to interact with the system.
  + Evaluate the system’s accuracy in answering real-world company policy related queries.

**Your daily GenAI Powered HR’s companion. An Expert guidance for your workplace needs.**

**By Manikandan Pandurangan (**[**Link to PPT**](https://app.presentations.ai/view/lbauCC)**) (**[**Link to Code**](https://github.com/Manisoft55-lab/Gen_AI_HR_Assistant.git)**)**

**Problem Statement: Your daily GenAI Powered HR query companion. An Expert guidance for your workplace needs.**

 Human Resource (HR) professionals often spend significant time searching for relevant company policies, responding to repetitive employee queries, and ensuring knowledge transfer across teams. Traditional document storage solutions make it difficult to retrieve specific policy information quickly, leading to inefficiencies and lost productivity.

This project aims to build an **AI-powered HR Knowledge Agent** that streamlines HR policy retrieval and query assistance using **LLM fine-tuning and Retrieval-Augmented Generation (RAG)**. By integrating with company document repositories like **Google Drive**, the system will:

1. **Reduce time spent by HR professionals** on searching policy documents.
2. **Improve response accuracy and consistency** for HR-related queries.
3. **Enhance HR productivity** by automating repetitive queries.
4. **Ensure up-to-date information retrieval** from dynamic company documents.
5. **Improve knowledge transfer** across HR teams without reliance on individual expertise.
6. **Utilize RAG for specific policy retrieval**, ensuring reliable and context-aware responses.

By implementing this solution, HR teams can focus on strategic tasks while employees receive instant, precise responses to their HR-related questions.

**Challenges in Building an AI-Powered HR Assistant**

1. **Integration with Google Drive Requires a Business Account** – Accessing corporate documents requires a **Microsoft Business Account** with strict authentication and permissions, making integration complex.
2. **Ensuring Accurate and Context-Aware Responses** – Fine-tuning the **LLM on HR-specific policies** and implementing **RAG** is essential to prevent **hallucinations** and provide factually correct responses.
3. **Maintaining Up-to-Date Policy Information** – HR policies change over time, requiring an **automated document sync** with Google Drive to ensure AI provides the latest information.

**Scope of the Project**

1. This HR Knowledge Agent can be integrated into a company’s existing ecosystem, allowing seamless retrieval of policy **documents** from various storage services such as **Google Drive** as of now. It currently supports **PDF files**, processes them using **text extraction and chunking techniques**, and enables efficient **retrieval-augmented generation (RAG)** for answering HR-related queries.

**Area of Improvement**

1. It can be developed as a standalone product that provides a flexible configuration interface to connect with different storage platforms. The system can easily adapt to varying company infrastructures, providing a **scalable solution** for HR departments. As the system evolves, it can support **multiple document storage services**, offering a unified approach to handling diverse data sources. This flexibility ensures that the HR assistant can be used across various organizations with minimal customization
2. An area of improvement for this project is to expand the **document format compatibility**, enabling the system to support a wider range of file types, including **Word**, **Excel**, **PowerPoint**, and **PDFs** from various storage services like **OneDrive, SharePoint**, **Confluence**, **NFS**, and others. Currently, the focus is on PDF files, but as organizations use diverse formats, enhancing support for these formats would significantly improve the system’s versatility.
3. This includes implementing **file format parsers** to extract text and relevant metadata from non-PDF files. Additionally, the ability to handle **scanned documents** through Optical Character Recognition (OCR) would improve the coverage of **image-based** or **handwritten policy documents**

**High Level Design**

