Project CompliAI: The Autonomous Regulatory Compliance Synthesizer

Abstract

India's rapid innovation is often slowed by the intricate and ever-changing landscape of regulatory compliance. Startups and Small and Medium-sized Enterprises (SMEs) face significant challenges in tracking, interpreting, and adhering to multi-jurisdictional laws, leading to financial penalties and stifled growth. This proposal outlines CompliAI, an autonomous AI agent designed to address this problem. CompliAI acts as a proactive, expert consultant that synthesizes and analyzes regulatory information from diverse sources. Using IBM's cutting-edge Generative AI technology, including the Granite series of models and the Agent Development Kit (ADK), our agent will provide structured compliance reports, identify non-obvious regulatory conflicts, and recommend actionable resolutions. This solution will dramatically reduce the burden on businesses, de-risking operations and accelerating innovation in India.

1. Solution Approach

Our solution, CompliAI, is an AI agent that automates the complex and resource-intensive task of regulatory research and compliance synthesis. The agent's end-to-end workflow is designed to be a transparent and autonomous process, going beyond simple document retrieval to perform high-level reasoning. **User Input & Semantic Deconstruction:** The user provides a natural language description of their business idea or product. Using the IBM Granite model, the agent intelligently deconstructs this input to identify key entities, business activities, and geographic jurisdictions. For example, a query about a "drone-based agricultural service in Maharashtra" would be understood in terms of drones, data sales, agriculture, and the specific state of Maharashtra. **Autonomous Multi-Source Research:** The agent, orchestrated by the IBM Agent Development Kit (ADK), autonomously calls a suite of custom-built tools to gather information from various sources. These tools are designed to access unstructured data from multiple government websites (e.g., DGCA, Ministry of Corporate Affairs, state-level portals), public APIs, and legal databases. **Dynamic Knowledge Graph Synthesis:** The core of our innovation lies in this step. Instead of simply collecting documents, the IBM Granite model processes all retrieved information to build an internal, dynamic "Knowledge Graph." This graph semantically links regulatory clauses and statutes, allowing the agent to understand dependencies, relationships, and, most importantly, potential conflicts between different laws. **Reasoning & Conflict Resolution:** The agent then analyzes the user's business proposition against its synthesized Knowledge Graph. It identifies direct obligations, highlights potential risks, and proactively pinpoints non-obvious contradictions. The agent's output is a structured report that not only lists compliance requirements but also provides clear, actionable recommendations to resolve identified conflicts. This demonstrates genuine AI reasoning and is the central "showstopper" of our solution. **Proactive Monitoring:** Once a business profile is established, the agent subscribes to its knowledge graph. If new policies or legal rulings are released that impact the business, it will automatically trigger an update and alert the user.

2. Tools, Libraries, and Datasets

Our technical stack is built around the core technologies specified by the hackathon, demonstrating our ability to leverage IBM's open-source assets for a real-world problem. **IBM Agent Development Kit (ADK):** This is the foundational framework for our project. We will use it to orchestrate the entire multi-step agent workflow, manage tool calls, and control the flow of information between our various components. **IBM Granite Series of Models:** We will utilize the Granite foundation models as the intelligent "brain" of our agent. Its advanced capabilities in natural language processing (NLP), reasoning, and text summarization are essential for deconstructing legal queries, building the Knowledge Graph, and synthesizing the final compliance report. **Open-Source Libraries:** To build our agent's external "tools," we will use widely-adopted Python libraries: - requests and BeautifulSoup for building a basic web scraping tool to pull data from government

websites. - pdfplumber for efficiently extracting text from PDF regulatory documents. - docker for local development and running the ADK server. **Datasets:** Given that no specific datasets are provided for this track, we will build a dynamic, custom dataset for our prototype. This dataset will be comprised of real, publicly available regulatory documents scraped from the official websites of Indian ministries, departments, and government bodies.

3. Expected Outcome

Our team plans to deliver a functional and compelling prototype of the CompliAI agent. Our submission will be judged on a live demonstration of the solution's core functionality, which includes: **A Functional, End-to-End Demo:** We will showcase a live walk-through where a user submits a business idea and the agent autonomously delivers a detailed, structured compliance report. **Proof of Concept for AI Reasoning:** We will specifically highlight the agent's ability to identify a complex regulatory conflict between multiple documents and provide a clear, recommended solution—a capability that is beyond the scope of traditional search engines or chatbots. **High-Impact Value Proposition:** CompliAI is a solution designed for the Indian context, directly addressing a critical pain point for millions of startups and SMEs. By simplifying legal compliance, our project will serve as a powerful proof-of-concept for how agentic AI can truly accelerate economic growth and innovation.