This presentation may include images from other products and services. These images are used for illustrative purposes only. Unless explicitly stated there is no implied endorsement or sponsorship of these products by Impetus. All copyrights and trademarks are property of their respective owners.

IMPETUS

AWS & Impetus | GenAI Hackathon

AI Governance & Compliance

Multi-Agent AI Governance System

Sebastian Diaz Gaviria

Note: All solutions submitted during the event will become the property of Impetus Technologies. Impetus reserves the right to use, develop, and extend the submitted ideas further. Participants will not retain ownership rights or control over the idea or its future use by Impetus.

Introduction

Team of One: Sebastian Diaz Gaviria . Role: Full-Stack GenAI Developer & Architect.

Email: diazgaviriasebastian01@gmail.com

Telegram @sebastiandg

Problem statement

Generative AI, while powerful, presents significant risks: it can leak sensitive data, generate harmful content, and operate without a clear audit trail.

Businesses need a reliable way to enforce security, compliance, and ethical policies in real-time to mitigate these legal.

This project implements a real-time, multi-agent governance layer for a bilingual financial assistant chatbot, analyzing both user prompts and LLM-generated responses.

Solution approach

Solution details: We have developed an intelligent governance shield that intercepts and analyzes every interaction. Our system uses five specialized, autonomous agents that work in concert to apply a multi-layered defense and compliance strategy.

Key features:

- Five specialized agents: Prompt Guard, Policy Enforcer, Output Auditor, Advisory Agent, and Audit Logger.
- Fully bilingual UI (English/Spanish) with dynamic LLM responses.
- Role-Based Access Control (RBAC) simulation.
- User feedback mechanism and performance caching for optimization.

Scope of scaling: The serverless architecture (Lambda, Bedrock, DynamoDB) allows for seamless horizontal and vertical scaling to handle fluctuating loads without manual intervention.

Tech-stack selection

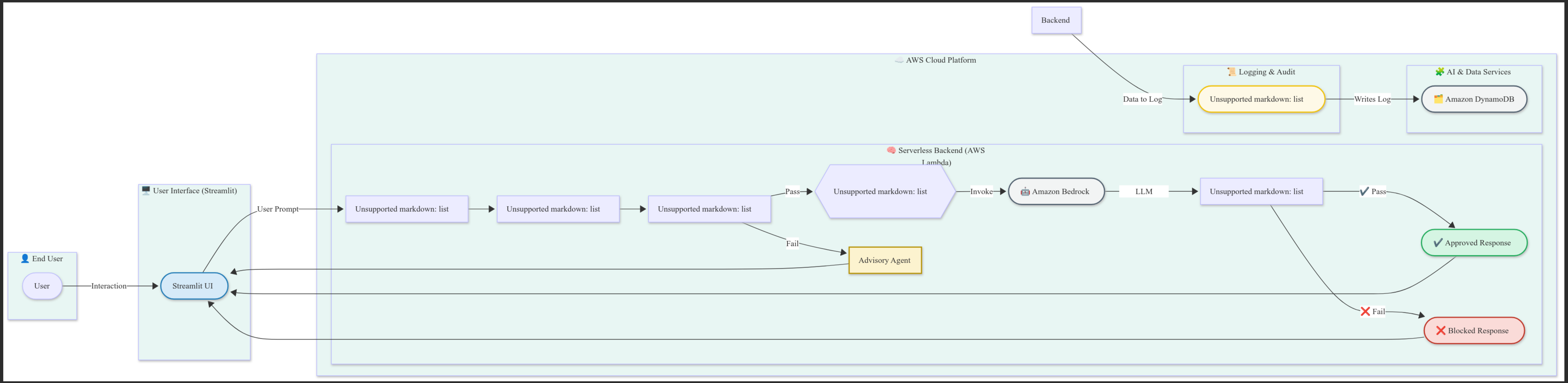
Selected model, language, framework: Cloude 3 Sonnet, Python, Streamlit

AWS services: Amazon Bedrock,Amazon Dynamo, AWS IAM

Give reasons for your tech stack selections :

- Streamlit was chosen for rapid, interactive UI development.
- Bedrock provides easy, scalable access to state-of-the-art foundation models.
- DynamoDB offers a highly scalable, low-latency, serverless database perfect for audit logs.

Architecture design/diagram



Cost estimates

Cost footprint: The chosen architecture is extremely cost-effective due to its serverless nature.

Hosting (Streamlit Cloud): $0 (Community Tier).

Database (DynamoDB): Effectively $0 for this scale, operating within the free tier.

The costs are minimal for a demonstration application (less than $10/month), for a considerable number of calls.

We are operating everything on the free tier of AWS.

Impact potential & limitations

Potential impact:

- Enables safe and responsible adoption of Generative AI across regulated industries.
- Reduces legal, financial, risks.
- Increases user trust and ensures compliance with regulations

Limitations of the current solution

- The governance ruleset is illustrative and would require expansion for a specific production environment.
- The Output Auditor's keyword-based check could be enhanced with a more sophisticated AI model for nuanced analysis.

Ethical & security considerations

Este proyecto está fundamentalmente diseñado para abordar problemas éticos y de seguridad. Por diseño, nuestro sistema:

- Prevents data leaks via the Policy Enforcer.
- Blocks harmful content with the Prompt Guard.
- Ensures transparency through the AI-powered Advisory Agent.
- Provides a full audit trail via the Audit Logger for accountability.

Project URL, video pitch & demo

<Include the hosted URL of the working project, if applicable>

<Include a short video (3 minutes) where you present your approach & architecture>

<This will help convey your solution approach and understanding of the project>

<Include a dry-run demo of the application>

<Include the link to your video (uploaded over a shareable drive)>

Thank you