Sentinel Integration Guide

AI Safety & Factuality Middleware

Built by GREM.xv

Overview

Sentinel acts as a middleware layer between your application and your LLM provider. It evaluates content for safety risks and factual reliability before output reaches end users.

Step 1 – Install Dependencies

```
pip install -r requirements.txt
```

Step 2 – Set Environment Variables

```
Linux / Mac:
export OPENAI_API_KEY=your_openai_key
export API_KEY=your_client_key

Windows (PowerShell):
setx OPENAI_API_KEY "your_openai_key"
setx API_KEY "your_client_key"
```

Step 3 – Run Sentinel

```
uvicorn main:app --host 0.0.0.0 --port 8000
```

Step 4 – Test Endpoint

```
curl -X POST http://localhost:8000/check   -H "Content-Type: application/json"   -H "Authorization:
```

Step 5 – Integrate Into Your Application

Send user input to /check endpoint. Read structured JSON response. Display safer_response or corrected_response to end user. Use safety_score and factuality_score for logging or moderation control.