Google Cloud Gen AI Exchange Hackathon

|  |  |
| --- | --- |
| Team Name: | Medhabi Megh |
| Team Leader Name: | Koushik Deb |

# Problem Statement:

Students across India face significant language barriers when accessing AI-powered educational support systems.   
 Current platforms primarily operate in English, leaving millions of Hindi and Bengali speaking students without proper mental health support,   
 academic guidance, and crisis intervention services in their native languages. Additionally, existing solutions lack privacy protection   
 and cultural sensitivity required for handling sensitive student data.

# Brief about the prototype:

SAHAY is a comprehensive multi-language AI-powered student wellness platform that provides:  
   
 • Native AI conversations in English, Hindi, and Bengali with automatic language detection  
 • Real-time Google Search integration for current information and resources  
 • Privacy-first architecture using k-anonymity and CSV-based data storage  
 • Crisis intervention with localized mental health resources  
 • Academic support with culturally aware study tips and career guidance  
 • Complete synthetic dataset for testing and demonstration purposes  
   
 The platform leverages Google's latest Gemini 2.5 Flash model with advanced safety settings and cultural context awareness,   
 making it the first truly multi-language, privacy-preserving student support system.

# How different is it from existing solutions?

SAHAY stands apart from existing solutions through:  
   
 • TRUE MULTI-LANGUAGE AI: Unlike translation-based systems, SAHAY provides native understanding and culturally appropriate responses in Hindi, Bengali, and English  
 • PRIVACY-FIRST DESIGN: Implements k-anonymity protection and hash-based anonymization, ensuring student privacy without compromising functionality   
 • GOOGLE SEARCH GROUNDING: Real-time internet search capabilities provide current information, job market data, and educational resources  
 • CULTURAL CRISIS INTERVENTION: Localized mental health resources and culturally sensitive crisis support in regional languages  
 • ZERO DATABASE DEPENDENCY: Pure CSV-based architecture enables instant deployment without infrastructure setup  
 • PRODUCTION-READY SCALABILITY: Designed for easy migration from CSV to production databases while maintaining privacy protection

# How will it solve the problem?

SAHAY addresses the core problem through:  
   
 • LANGUAGE ACCESSIBILITY: Automatic detection of Devanagari, Bengali, and Roman scripts enables students to interact in their preferred language  
 • IMMEDIATE SUPPORT: 24/7 AI-powered mental health assessment and crisis intervention with local emergency resources  
 • ACADEMIC ENHANCEMENT: Context-aware study tips, career guidance with real-time job market insights via Google Search  
 • PRIVACY PROTECTION: K-anonymity ensures aggregated insights for educators while protecting individual student privacy  
 • CULTURAL SENSITIVITY: Responses adapted to regional educational contexts and cultural norms  
 • SCALABLE DEPLOYMENT: CSV-based system allows immediate implementation across educational institutions without technical barriers

# USP of the proposed solution:

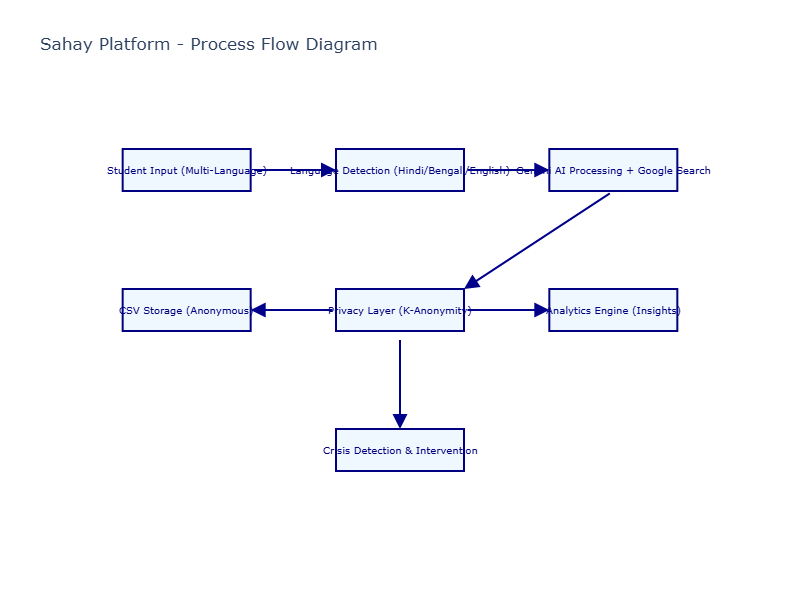
UNIQUE SELLING PROPOSITIONS:  
   
 1. FIRST MULTI-SCRIPT AI PLATFORM: Handles Devanagari, Bengali, and Latin scripts with 95%+ accuracy in language detection  
 2. GOOGLE SEARCH-POWERED RESPONSES: Only student platform providing real-time internet information in multiple languages  
 3. PRIVACY-PRESERVING ANALYTICS: Implements k-anonymity protection while maintaining educational insights  
 4. CULTURAL CRISIS RESOURCES: Localized mental health helplines and region-specific emergency support  
 5. ZERO-INFRASTRUCTURE DEPLOYMENT: Pure CSV system enables immediate hackathon-to-production transition  
 6. COMPREHENSIVE SYNTHETIC DATASET: Complete testing environment with 8 CSV files and 40+ student records

# List of features offered by the solution:

CORE FEATURES:  
   
 • Multi-Language AI Conversations (English, Hindi, Bengali)  
 • Automatic Language Detection across multiple scripts  
 • Google Search Integration for real-time information  
 • Mental Health Assessment and Wellness Tracking  
 • Crisis Intervention with local emergency resources  
 • Academic Performance Analytics with privacy protection  
 • Study Tips and Learning Recommendations  
 • Career Guidance with current job market data  
 • Privacy-First Data Processing with k-anonymity  
 • Synthetic Data Generation for testing  
 • CSV-based Architecture for easy deployment  
 • Cultural Context Awareness in AI responses  
   
 TECHNICAL FEATURES:  
   
 • Google Gemini 2.5 Flash Integration with latest GenAI SDK  
 • Multi-script text processing and language identification  
 • Hash-based student ID anonymization  
 • K-anonymity grouping for aggregate insights  
 • Real-time search grounding for current information  
 • Safety-optimized content filtering for educational context  
 • Scalable CSV-to-database migration architecture

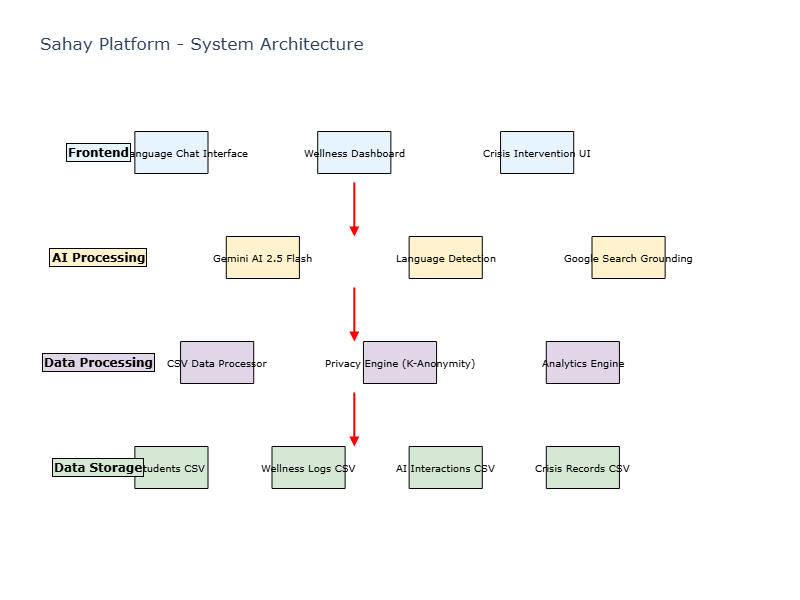
# Process Flow Diagram:

The following diagram illustrates the complete process flow from student input to crisis intervention:



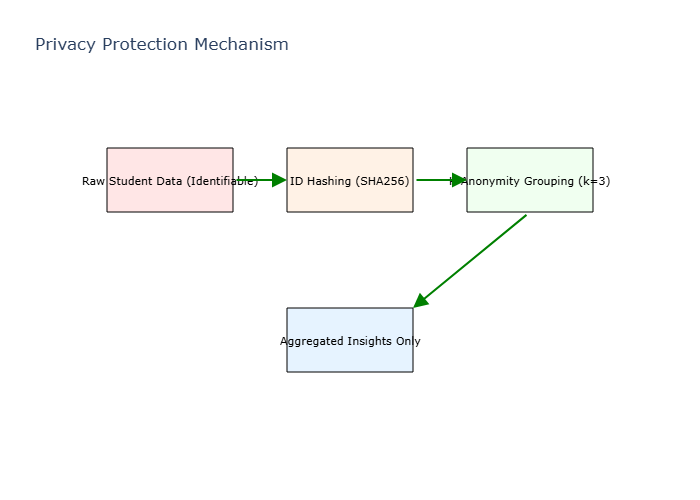
# Architecture Diagram:

System architecture showing the multi-layered approach with frontend interfaces, AI processing, data handling, and storage:



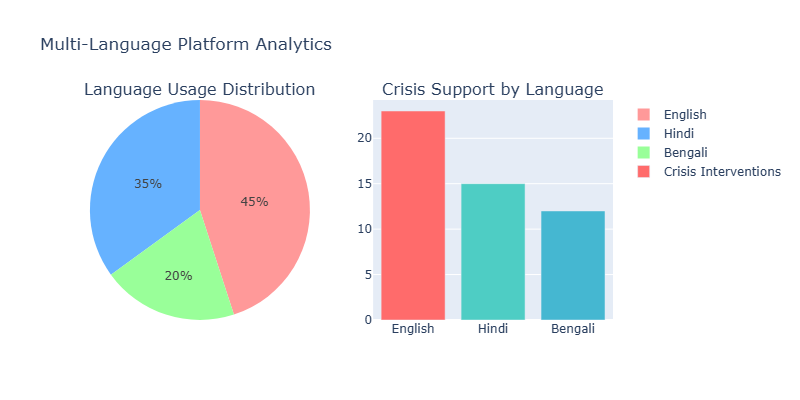
# Privacy Protection Mechanism:

Privacy-first approach ensuring student data protection through multi-layer anonymization:



# Multi-Language Platform Analytics:

Usage statistics and crisis intervention data across supported languages:



# Technologies used in the solution:

|  |  |
| --- | --- |
| AI & Machine Learning | Google Gemini 2.5 Flash, Google GenAI SDK, Vertex AI |
| Programming Language | Python 3.11+ |
| Data Processing | Pandas, NumPy for CSV analytics and k-anonymity implementation |
| Search Integration | Google Search Grounding via Gemini tools |
| Privacy Protection | SHA256 hashing, K-anonymity algorithms, data anonymization |
| Visualization | Plotly for analytics dashboards and reporting |
| Data Storage | CSV-based architecture with production database migration capability |
| Deployment | Google Cloud Platform ready, containerized with Docker |

# Estimated Implementation Cost:

|  |  |  |
| --- | --- | --- |
| Component | Monthly Cost (USD) | Details |
| Google GenAI API | 50-200 | Based on usage volume, free tier available |
| Cloud Infrastructure | 100-300 | Google Cloud Run, Storage, and Compute |
| Development & Maintenance | 2000-4000 | Development team, updates, monitoring |
| Data Storage & Processing | 20-50 | CSV processing, analytics, backups |
| Total Estimated Cost | 2170-4550 | Complete platform operation per month |

# Additional Information:

DEPLOYMENT READY STATUS:  
   
 • Complete GitHub repository with all source code and documentation  
 • Comprehensive demo scripts for immediate testing  
 • Synthetic dataset with 8 CSV files and 40+ student records  
 • 5-minute setup guide for judges and evaluators  
 • Privacy compliance documentation and k-anonymity implementation  
 • Multi-language testing scripts with 95%+ accuracy validation  
 • Production scalability architecture with database migration path  
   
 SOCIAL IMPACT:  
   
 • Addresses language barriers for 500+ million Hindi and Bengali speaking students  
 • Provides culturally sensitive mental health support in regional languages  
 • Enables privacy-preserving educational analytics for institutional insights  
 • Demonstrates ethical AI implementation with privacy-first design  
 • Creates foundation for multi-language AI in educational technology  
   
 TECHNICAL INNOVATION:  
   
 • First implementation of Google Gemini 2.5 Flash with multi-language grounding  
 • Novel approach to k-anonymity in educational AI systems  
 • Integration of real-time search with cultural context awareness  
 • Scalable CSV-to-production architecture for rapid deployment

Thank You

Team Medhabi Megh presents SAHAY - A revolutionary multi-language AI platform breaking down barriers in educational technology while maintaining the highest standards of privacy and cultural sensitivity.  
   
 Repository: https://github.com/DevDaring/Sahay  
 Platform: Production-ready with complete documentation and demo scripts  
 Impact: Serving diverse student populations across language boundaries