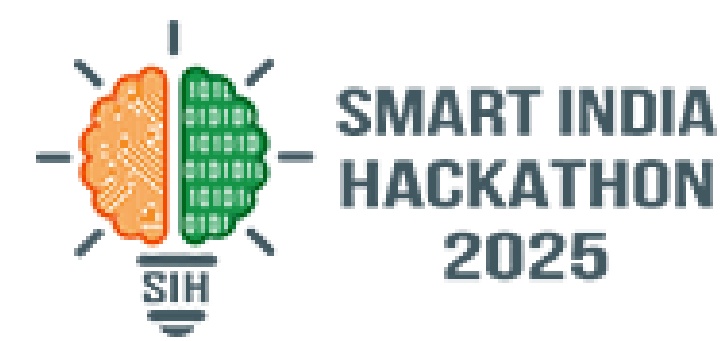


# SMART INDIA HACKATHON 2025



## TITLE PAGE

- Problem Statement ID – SIH25066
- Problem Statement Title- AI-driven ChatBOT for INGRES as a Virtual Assistant (Ministry of Jal Shakti, Software): Matches Digital India service delivery and Jal Shakti's push for AI decision tools, with precedent from agri-scheme chatbots.
- Theme- Smart Automation
- PS Category- Software
- Team ID-
- Team Name (Registered on portal) - JALBUDDY



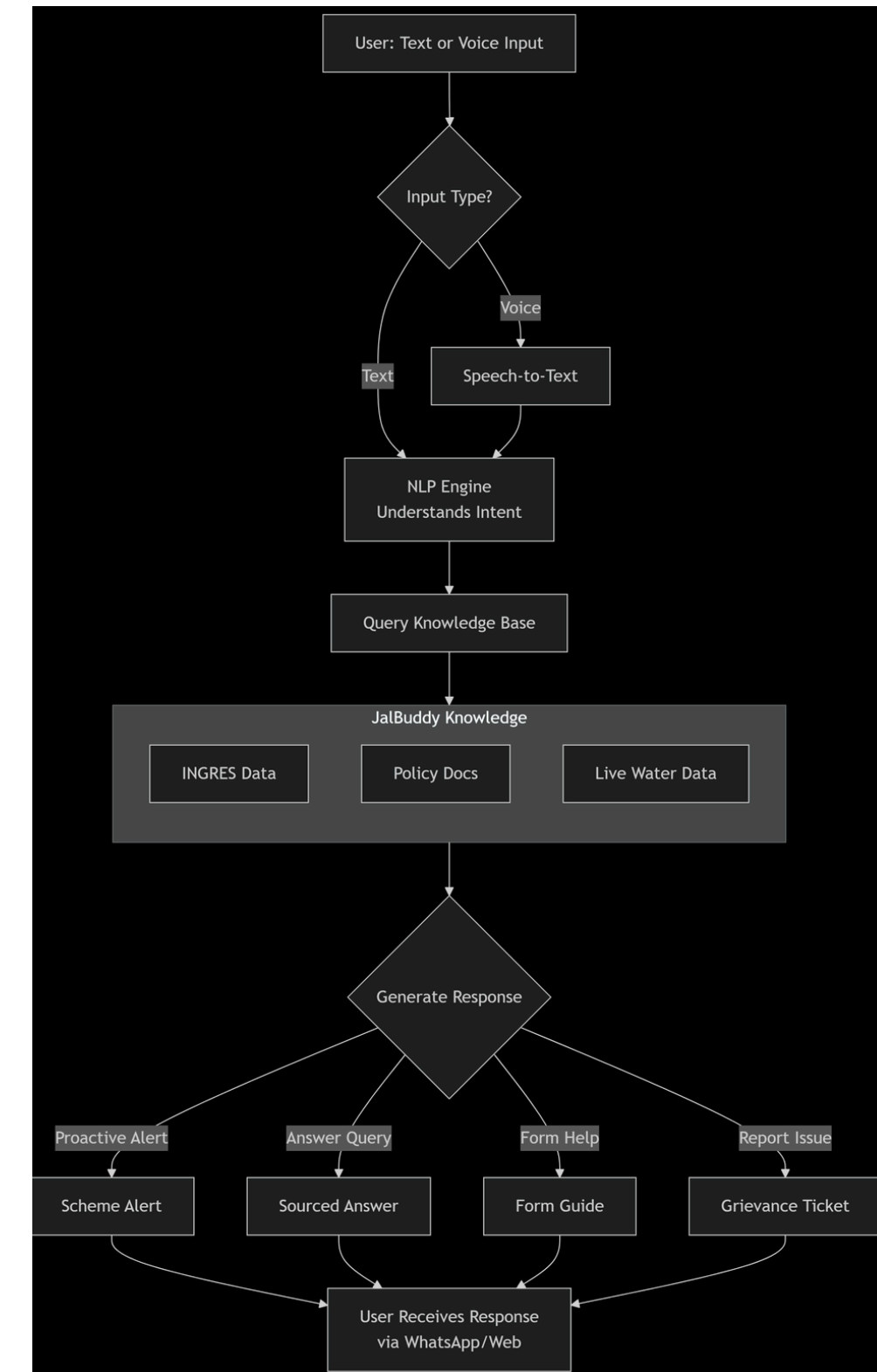
# JalBuddy

## • Proposed Solution

- An AI-powered conversational assistant that transforms complex groundwater data access through natural language processing, enabling farmers, officials, and citizens to query INGRES database using simple voice/text commands in multiple languages.
- Detailed explanation of the proposed solution:
- Conversational AI Interface: Natural language processing converts user queries into precise INGRES data retrievals with GEC-2015 methodology explanations
- Multi-Channel Accessibility: WhatsApp Bot, Progressive Web App, USSD gateway, and voice interfaces for comprehensive rural reach
- Hybrid Online-Offline Architecture: Core functionality works without internet, syncs when connected for rural accessibility
- Real-Time Intelligence: Model Context Protocol integration provides live DWLR telemetry, weather data, and groundwater alerts

## • Innovation and uniqueness

- Proactive Alerts: Doesn't just answer questions; analyzes data to send users timely alerts about water schemes, conservation tips, and drought warnings.
- First voice-enabled INGRES assistant with offline capabilities for rural areas
- AI-powered GIS query translation from natural language to spatial data visualization
- Multi-modal accessibility design supporting illiterate users through voice interaction
- Predictive analytics using DWLR data for seasonal water availability forecasting
- Integrate with the Bhuvan map platform or a simple GIS interface.
- For every response it cites its source provides a link to the PDF. It can also show a "confidence score."
- provide form guidance for people with lack of understanding

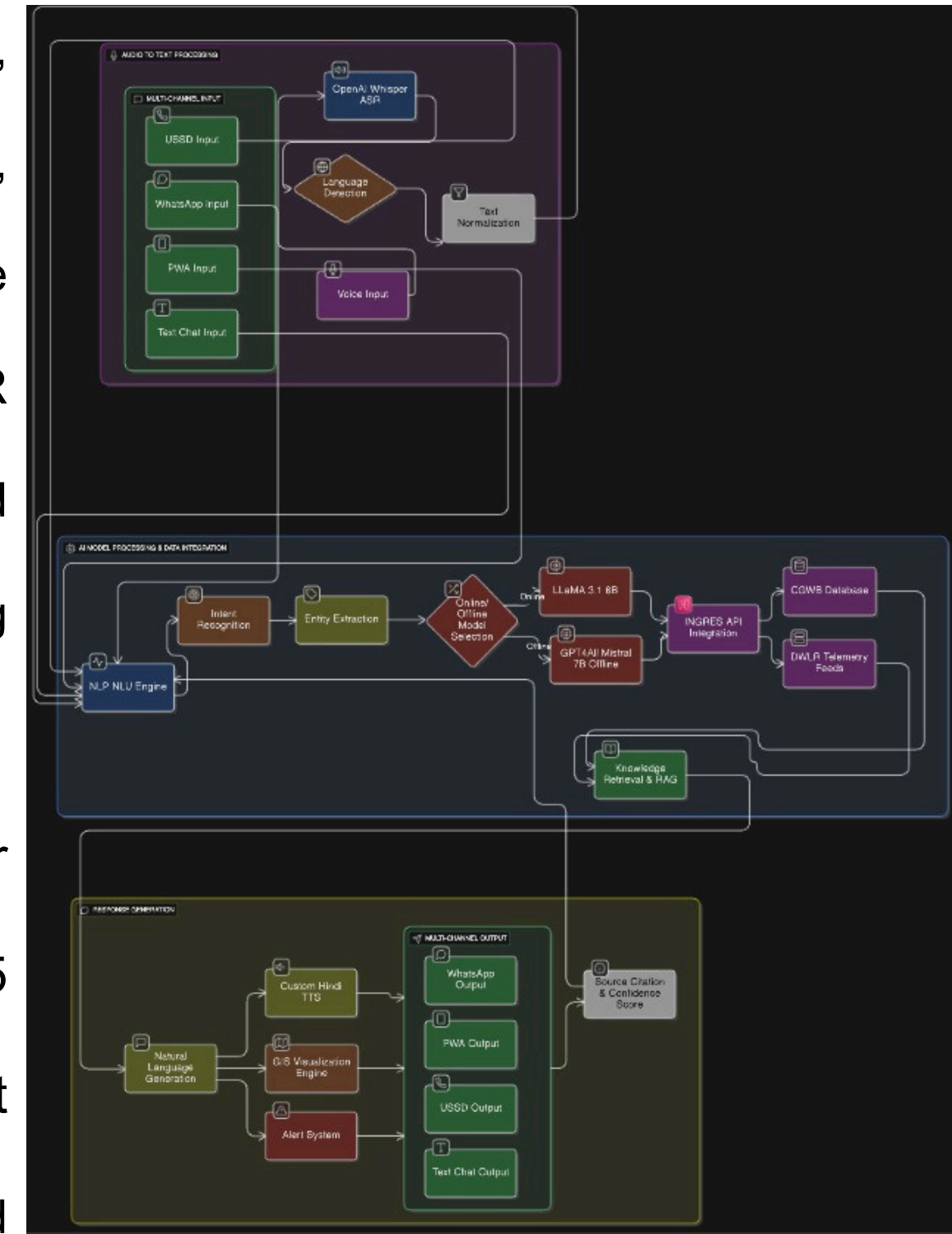


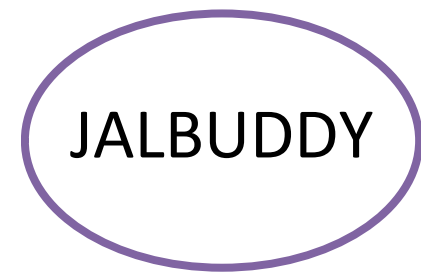
## • Technologies to be used

- AI/ML Stack: Fine-tuned LLaMA 3.1 8B, GPT4All Mistral 7B (offline), OpenAI Whisper ASR, Custom Hindi TTS
- Backend Services: Spring Boot microservices, Redis caching, Apache Kafka streaming, PostgreSQL database
- Frontend Interfaces: React PWA, WhatsApp Business API, USSD gateway integration, Voice UI framework
- Data Integration: Model Context Protocol (MCP), INGRES APIs, CGWB databases, DWLR telemetry feeds
- Infrastructure: Docker containerization, Kubernetes orchestration, AWS/Azure cloud deployment
- Security & Compliance: End-to-end encryption, CERT-In compliance framework, audit logging systems

## • Methodology

- Phase 1: RAG system development with INGRES API integration and groundwater domain fine-tuning
- Phase 2: Multi-channel interface deployment with basic query resolution and GEC-2015 explanations
- Phase 3: Offline mode implementation with DWLR monitoring and predictive alert systems
- Phase 4: Field testing across diverse geographical regions with iterative refinement and optimizations.





# FEASIBILITY AND VIABILITY



- **FEASIBILITY AND VIABILITY**

- Technical Feasibility: INGRES provides public APIs with comprehensive documentation and real-time data access
- Resource Availability: Open-source AI models and government data sources minimize development barriers
- Infrastructure Scalability: Cloud-native architecture supports millions of concurrent users with auto-scaling capabilities
- Government Integration: Established CGWB partnerships and Ministry of Jal Shakti AI initiatives provide institutional support

- **Potential challenges and risks**

- Data Consistency: update frequencies across different INGRES and regional databases
- Rural Connectivity: Limited internet access in remote areas affecting real-time data synchronization
- User Adoption: Digital literacy among target rural people requiring extensive training
- Language Complexity: Technical groundwater terminology translation accuracy across multiple regional languages

- **Strategies for overcoming these challenges**

- Hybrid Architecture: 80% functionality available offline with intelligent sync when connectivity restored
- Multi-Modal Interface: Voice-first design with visual aids accommodating varying literacy levels
- Community Deployment: Partnership with Panchayati Raj institutions for grassroots adoption and training
- Incremental Rollout: Phased deployment starting with high-connectivity districts, expanding based on success metrics
- Continuous Learning: User feedback integration for improving language models and query accuracy



- **Potential impact on the target audience**

- Farmers & Rural Communities: Access to groundwater data reduces failed drilling attempts by 70%, saving ₹50,000+ per household
- Government Officials: 60% reduction in helpdesk queries with automated information delivery and policy compliance checking
- Water Resource Managers: Real-time monitoring enables proactive management of 1,186 over-exploited assessment units
- Digital Inclusion: Voice-first design enables participation by 850+ million users with low digital literacy

- **Benefits of the solution**

- **Economic Benefits:**

- ₹2,000+ crore annual savings from prevented unsuccessful drilling attempts nationwide
- 15-20% agricultural productivity improvement through optimized groundwater utilization
- Reduced administrative costs for government information dissemination and query resolution

- **Social Benefits:**

- Democratized access to technical groundwater information for marginalized communities
- Enhanced water security for rural households through predictive availability alerts
- Improved disaster preparedness with early warning systems for drought-prone regions

- **Environmental Benefits:**

- Prevention of over-extraction in critical/semi-critical groundwater areas
- Data-driven conservation promoting sustainable aquifer management
- Reduced environmental degradation through informed groundwater usage decisions

## • Details / Links of the reference and research work

- INGRES Platform: <https://ingres.iith.ac.in>
- Central Ground Water Board (CGWB): <https://cgwb.gov.in/>
- GEC-2015: [https://cgwb.gov.in/sites/default/files/inline-files/2020-1117\\_Detailed\\_Guidelines\\_GEC2015.pdf](https://cgwb.gov.in/sites/default/files/inline-files/2020-1117_Detailed_Guidelines_GEC2015.pdf)
- India-WRIS: <https://indiawris.gov.in/wris/>
- Whatsapp business API: <https://developers.facebook.com/docs/whatsapp/>
- Model Context Protocol (MCP): <https://modelcontextprotocol.io>
- OpenAI Whisper: <https://openai.com/research/whisper>
- HuggingFace Transformers: <https://huggingface.co/docs/transformers/index>
- GPT4All Offline Models: <https://www.nomic.ai/gpt4all>