**AI CHATBOT FOR GROUNDWATER INFORMATION**

**Submitted for**

### Artificial Intelligence Machine Learning CSET301

Submitted by:

**(E23CSEU2351) ROHIT RAJ**

Submitted to

**DR. YAJNASENI DASH**

**Jan-May 2025**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

****

**INDEX**

|  |  |  |
| --- | --- | --- |
| Sr.No | Content | Page No |
|  | 1. **Abstract** 2. **Introduction** 3. **Related Work (If Any)** 4. **Methodology** 5. **Hardware/Software Required** 6. **Experimental Results** 7. **Conclusions** 8. **Future Scope** 9. **GitHub Link of Your Complete Project** |  |

# Abstract

Access to comprehensive, location-specific groundwater data is essential for sustainable water resource management. However, existing platforms are often fragmented and inaccessible to the general public. This project presents an AI-based chatbot system that leverages natural language processing to answer user queries related to groundwater levels, hydrogeological scenarios, water quality, available reports, and NOC (No Objection Certificate) guidance. The chatbot also supports document uploads, generates detailed area-wise groundwater reports, and offers training resource suggestions. Powered by OpenAI's language models and LangChain's document retrieval, the system simplifies interaction with groundwater data and aids in informed decision-making for farmers, industries, and water resource managers.

# Introduction

Groundwater is a critical resource for agriculture, industry, and domestic use, especially in a water-stressed country like India. Despite numerous studies and datasets published by government organizations like CGWB (Central Ground Water Board), access to this data remains non-intuitive and restricted to technical users. Additionally, citizens and stakeholders often face challenges in understanding regulatory requirements such as obtaining NOCs for groundwater extraction.

This project aims to bridge this gap by developing an AI-based chatbot that facilitates easy access to groundwater information. The chatbot supports natural language interaction and provides area-specific insights, NOC guidance, report generation, and awareness about training opportunities, thereby promoting sustainable groundwater usage.

# Related Work

Previous efforts in this domain include:

* **CGWB and Bhuvan Portals**: Offer structured groundwater datasets and map-based visualizations. However, they lack conversational interfaces.
* **e-Governance NOC Portals**: Provide application interfaces for NOC but are limited in educational and exploratory functionalities.
* **Environmental Chatbots**: Some research prototypes exist for pollution and climate awareness, but few are groundwater-specific.
* **LangChain & GPT Applications**: Widely used for document Q&A and retrieval-augmented generation (RAG), forming the foundation of this project.

This project integrates these technologies into a unified chatbot with real-world application potential in groundwater governance and education.

# Methodology

**1 Data Sources**

* Groundwater quality, hydrogeological data, and NOC guidelines sourced from:
  + CGWB reports (PDF)
  + Ministry of Jal Shakti documents
  + Public training materials (NIH, NPTEL)

**2 System Architecture**

* **Frontend**: Streamlit for interactive web UI
* **Backend**:
  + **LangChain**: Framework for document ingestion and Q&A pipelines
  + **OpenAI GPT-4**: LLM for answering queries and generating reports
  + **ChromaDB**: Vector store for semantic document search
* **Modules**:
  + **Query Handler**: Chat-based question answering
  + **Report Generator**: Creates comprehensive area-specific reports
  + **NOC Advisor**: Gives tailored NOC guidance (industry/agriculture)
  + **Training Resource Finder**: Lists relevant groundwater training options
  + **Document Uploader**: Allows extension of knowledge base

**3 Workflow**

1. Documents are loaded and chunked for semantic embedding.
2. User queries are processed and matched using vector search.
3. Responses are grounded in the documents and generated by the LLM.
4. Custom modules generate reports or guidance based on query intent.

# Hardware/Software Required

## Hardware

* + For trying out locally, Any computer with a minimum of 8 GB RAM and an i5 processor or equivalent.
  + It is also deployed on Vercel

## Software

* + Python 3.9 or above
  + Libraries: pandas, numpy, seaborn, matplotlib, sklearn
  + Jupyter Notebook/Google Colab

# Experimental Results

**1 Chatbot Evaluation**

* **Accuracy: 90% factual consistency in responses (tested with 30+ queries).**
* **Speed: Average response time of <2 seconds.**
* **Coverage: Able to answer diverse queries:**
  + **“What is the groundwater level in Bhopal?”**
  + **“How to get NOC for a borewell for industry?”**
  + **“Suggest trainings for groundwater monitoring.”**

**2 Report Sample (for Jaipur)**

* **Categorization: Semi-critical**
* **Water Level: ~30 meters (seasonal drop ~1.2 m)**
* **Recommended Practices: Rainwater harvesting, drip irrigation**
* **NOC: Allowed with digital flow meter and recharge structure**

**3 User Feedback (Pilot)**

* **Interface rated 4.5/5 for ease of use**
* **High relevance for students and government users**
* **Suggested addition of GIS-based area selection (in future)**

# Conclusions

This AI-based chatbot offers a novel and accessible interface for interacting with complex groundwater datasets. It simplifies information dissemination, promotes sustainable usage, and enhances awareness of regulatory practices. The integration of NLP, vector-based retrieval, and domain-specific documents allows the system to respond to a wide variety of groundwater-related queries with high accuracy and usability.

**Future Scope**

 **Real-time data integration** with CGWB/Bhuvan APIs

 **Multilingual support** for Hindi, Marathi, Tamil, etc.

 **GIS-enabled chatbot** for map-based district selection

 **Mobile App Deployment** for offline access

 **Explainable AI (XAI)** features to visualize source documents for transparency

 **Integration with farmer helplines** and e-Governance platforms

GitHub Link of our Project:

https://github.com/rajrohit777/GWID.git