Team Name: **JugJugGeo**

Team Leader Name: **Divyanshu Pabia**

Problem Statement: 4

Designing a Chain-of-Thought-Based LLM System for Solving Complex Spatial Analysis Tasks Through Intelligent Geoprocessing Orchestration

Team Members

Team Leader

Name: Divyanshu Pabia
College: SRM Institute of Science and Technology, Kattankulathur, Chennai

Team Member-1

Name: Hridya Munjal
College: SRM Institute of Science and Technology, Kattankulathur, Chennai

Team Member-2

Name: Akshat Jain
College: SRM Institute of Science and Technology, Kattankulathur, Chennai

Team Member-3

Name: Krishna Shrivastava
College: SRM Institute of Science and Technology, Kattankulathur, Chennai

Problem Statement:

Across multiple domains in geospatial analysis, analysts and researchers must **integrate data** from multiple heterogenous sources and dozens of **GIS algorithms** just to answer a single question such as “Where will Chennai flood next monsoon?”. These workflows demand costly software, deep geospatial know-how, and days or weeks of manual scripting. 90% of the time in geospatial analysis is spent on data preparation and tool configuration. As a result, most organizations either outsource the work or rely on outdated heuristics, leading to slow decisions, opaque methods, and missed opportunities. Existing AI solutions provide “black box” results without explanation, making it difficult to verify results, understand assumptions or learn from the process. **Result - Critical spatial decisions are either delayed, expensive or made without proper analysis**.

Lack of GIS Expertise and Tool Knowledge ***Time inefficiency in data preparation and env setup***
Unverifiable AI results and hallucinations

Our Solution:

JugJugGeo collapses that complexity into a **natural-language chat**. A **domain-tuned** Large Language Model plans each analytic step in plain English, retrieves the right documents from a **vector knowledge base**, executes PyQGIS and WhiteboxTools code in a **sandbox**, pulls fresh Sentinel-2 or Bhoonidhi imagery on demand and STAC APIs, and streams back maps, statistics, and fully **cited Chain-of-Thought process** to the user. Because the engine is open-source, containerized, and already running on **AWS Infrastructure**, teams can self-host, plug in their own data, and start generating enterprise-grade spatial insight in seconds instead of weeks.

Ask and Analyze ***Transparent AI Reasoning*** ***Professional GIS Tools***
Personal Document Library ***Sandboxed Code Execution*** ***Web Search Integration***

Why JugJugGeo is a leap, not a tweak.

Pain Points	Existing Solutions & Their Limitations	How JugJugGeo Solves It
🎯 GIS Expertise Barriers	**GeoForge** – fine-tuned GPT-4 but locked to Swiss servers, requires technical setup **CARTO AI** – enterprise-only, expensive licenses	**Natural-language interface**: zero GIS training – ask in plain English and get professional analysis instantly
🔍 Lack of Transparency	**MapAI, EarthGPT** – black-box with no reasoning Commercial GPTs give generic answers without methodology	**Chain-of-Thought reasoning**: shows every step (data fetch → algorithm → parameters), building trust and enabling learning
🛠️ Tool Fragmentation	**GeoGPT**– limited fixed tool pool (custom) **Kue AI** – QGIS-only plugin **ChatGeoAI** – 25 % first-try success rate	**Complete integration**: PyQGIS (300+ algorithms) + WhiteboxTools (518+) + Planetary Computer + Bhoonidhi ISRO – all in one chat
📊 Limited Specialized Knowledge	**Commercial ChatGPT plugins** – general AI with shallow geo skills **Remote Sensing ChatGPT** – limited to image interpretation	**Domain-tuned models**: GeoEmbedding-7B + GeoReranker-568M trained on 500 K+ geospatial docs – understands spatial concepts like an expert
💰 High Cost & Complexity	Enterprise GIS licenses run **$1 k–10 k**+ per year Consultants add weeks & high fees	**Open-source deployment**: self-host or one-click AWS stack Turns weeks of work into seconds of conversation


Natural Language Interface
• Ask in Plain English
• No GIS Training Required
• Advanced Visualizations
• Live Health Status Monitoring
Deep Discovery Engine
• Multi-Step Research
• Cross-Source Analysis
• Comprehensive Reports
Chain-of-Thought Reasoning
• Transparent Process
• Explainable Decisions
Geo
Secure Code Execution
• Sandboxed Environment
• Interactive Results
GIS Tools Integration
• PyQGIS
• WhiteboxTools
Knowledge Base Management
• Document Upload
• Intelligent Indexing
Global Data Access
• Planetary Computer
• Bhoonidhi ISRO

Feature: Knowledge Base Management (Document Library)

"Transform Documents into Searchable Spatial Intelligence"

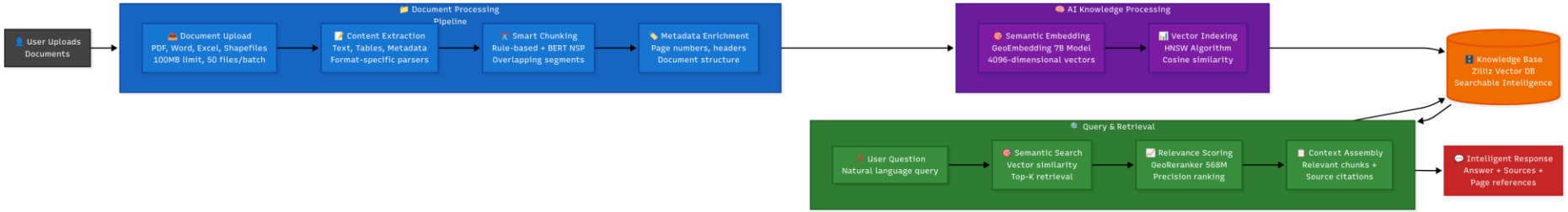
• **Bulk upload:** Drag-and-drop up to 50 files (≤ 100 MB each, PDF/Word/Excel/GeoJSON/Shapefiles) per batch.

• **Auto-parse & vectorise:** Every page is extracted, split, embedded with our **GeoEmbedding model** and indexed in Zilliz Cloud (4096-dim vector).

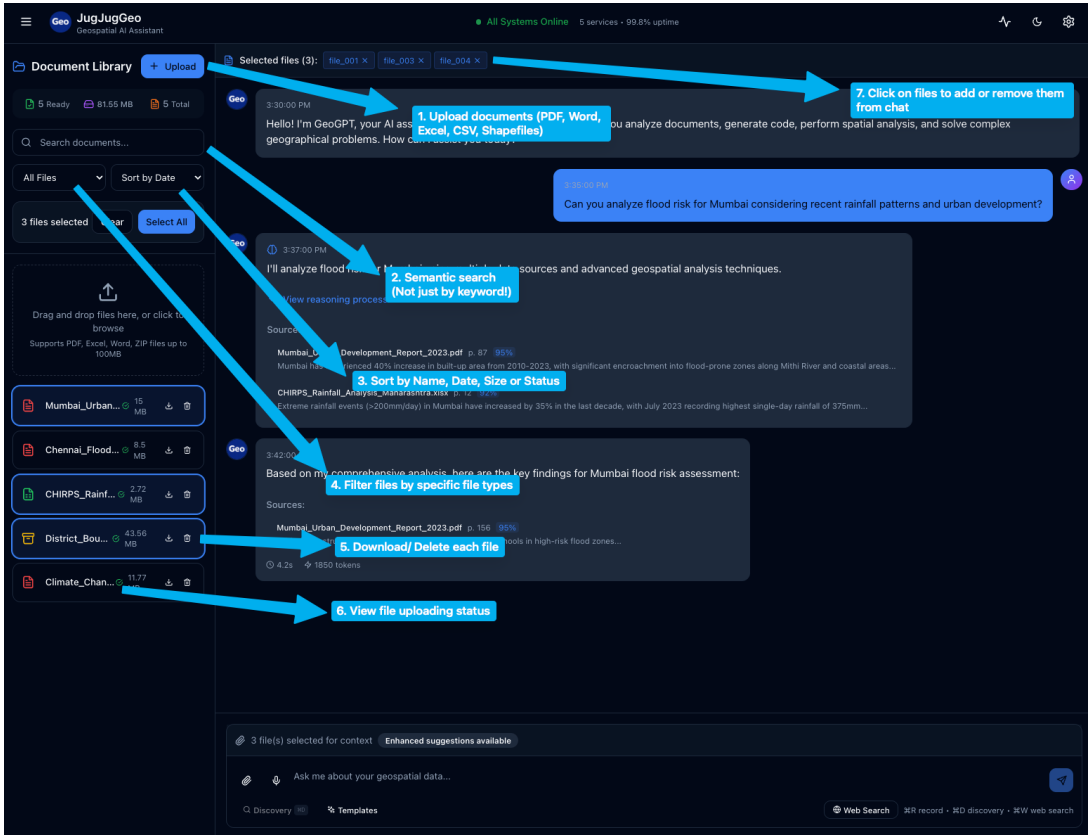
• **Semantic search that cites pages: GeoReranker** pinpoints the exact chunks. answers arrive with page-level citations, so evidence is always one click away.

• **Flexible library controls:** Folder, rename, move, delete, or download any document.

• **Ask once, learn forever:** Query the whole library or a single folder.



Feature: Knowledge Base Management (Document Library UI)



Feature: Chain of Thought Reasoning

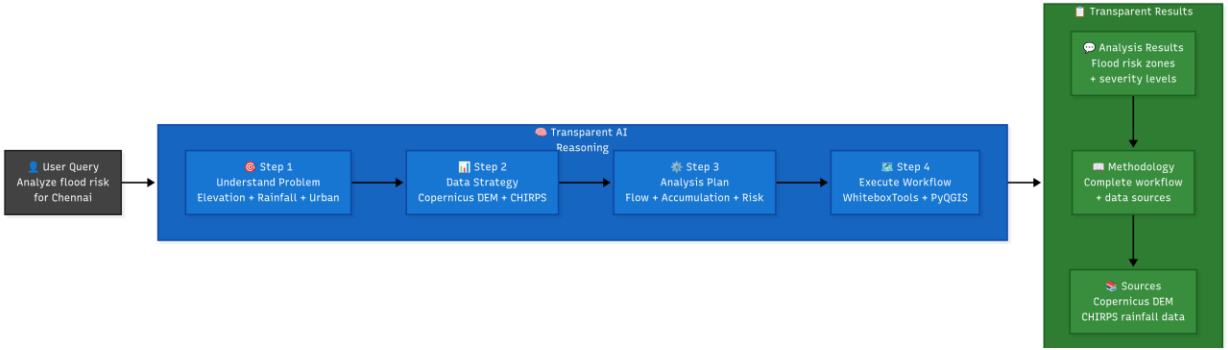
"Watch AI Think Like a Geospatial Expert"

• **Shows its work, not just the answer:** Each chat response arrives with a full "thinking" log: what data sets were chosen, which GIS tools are invoked, and why.

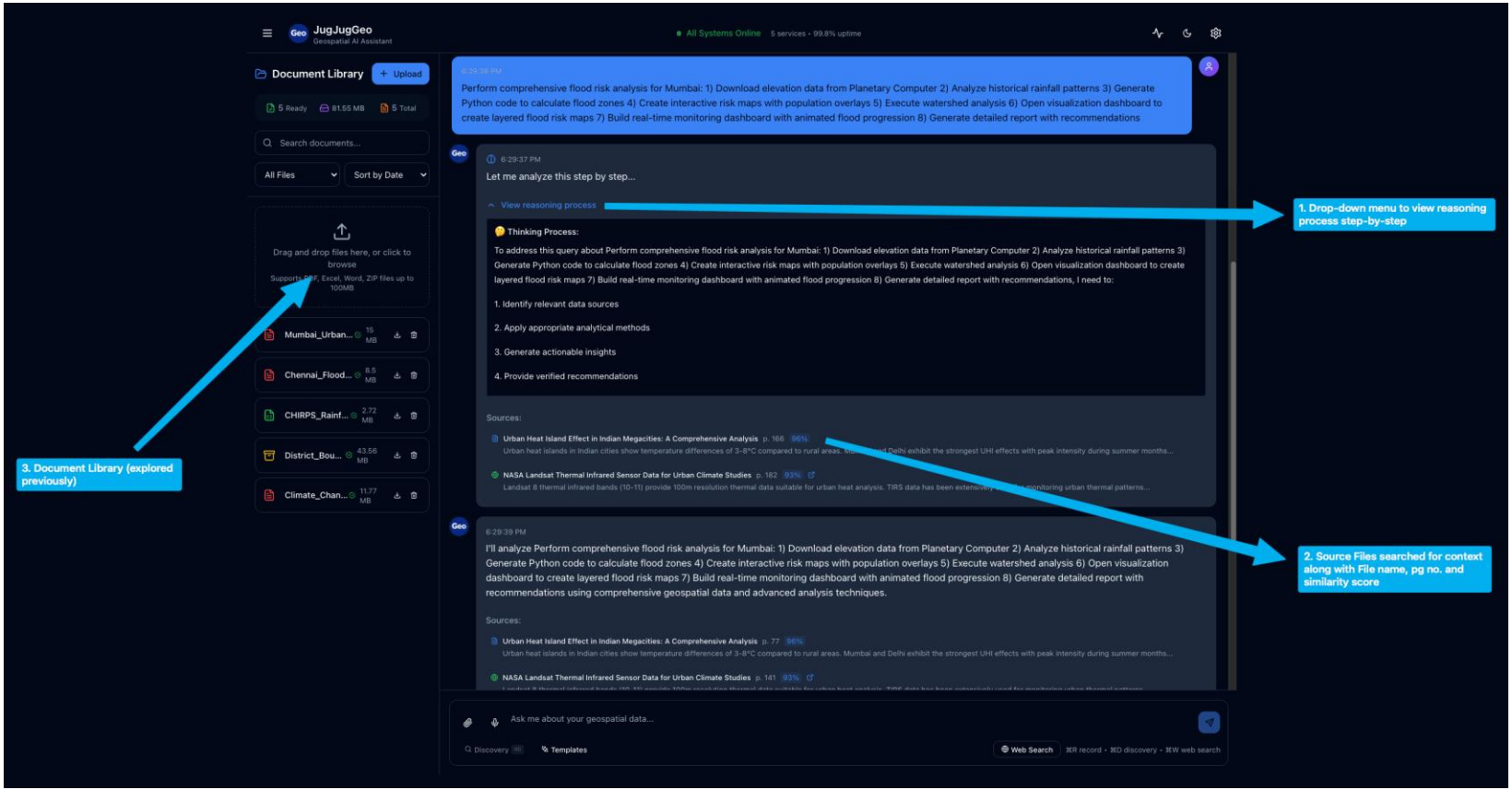
• **Expert-level decomposition:** GeoGPT breaks one query into numbered sub-tasks.

• **Verifiable & auditable:** Every step's rationale is stored with the conversation, so teams can review, reproduce, and hand it to regulators or peer reviewers.

• **Model tuned for reasoning:** GeoGPT-R1's 8 k-token window and specialized training optimized it for long, multi-step chains without hallucinating spatial facts.



Feature: Chain of Thought Reasoning (UI)



Feature: Deep Discovery Engine

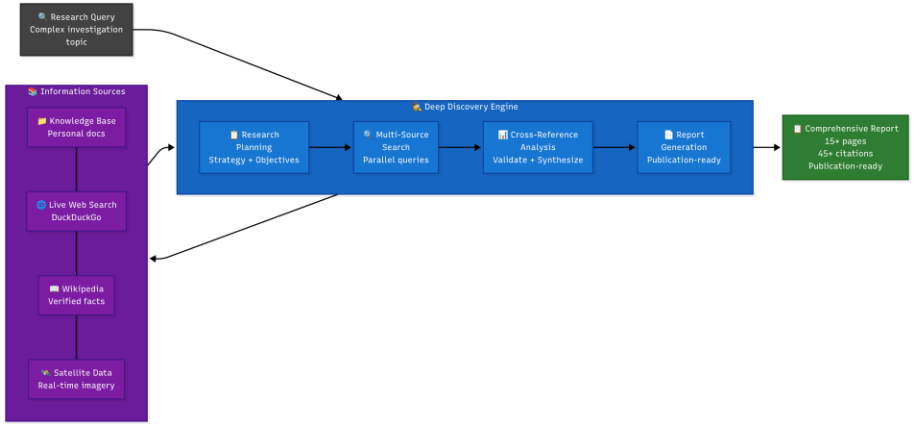
“Autonomous Multi-Step Geospatial Research”

• **Multi-step research orchestrator:** A single query triggers an automated **five-stage** workflow - scans web, Wikipedia, and your knowledge base, clusters ideas, and homes in on high-value sources.

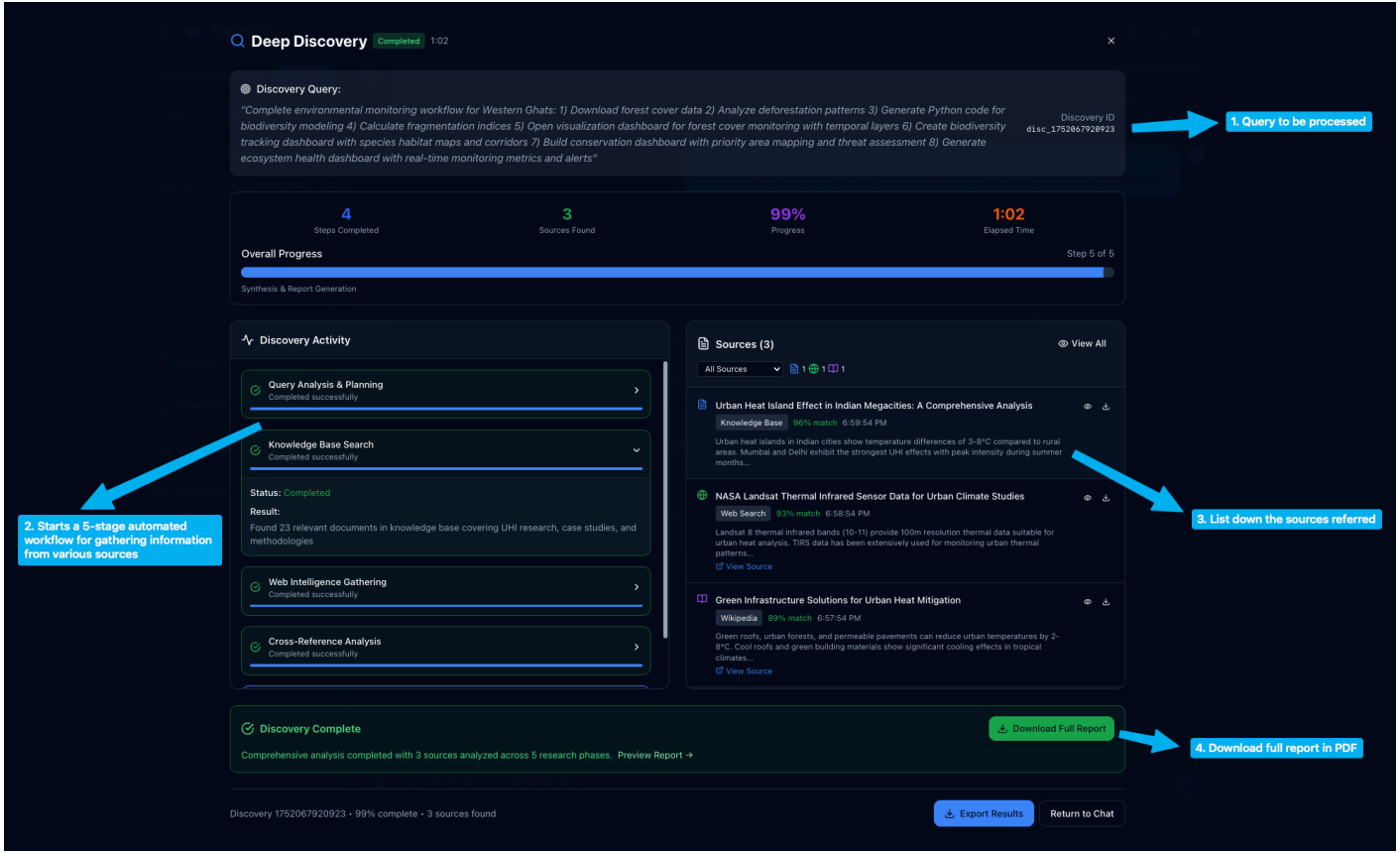
• **Live ‘Activity + Sources’ pane:** Watch the engine read, reason, and cite in real-time, every URL, document chunk, and relevance score is exposed for audit.

• **Configurable depth & focus:** Choose up to max_steps = 10, toggle web vs KB search - all via the **/discovery/start** endpoint.

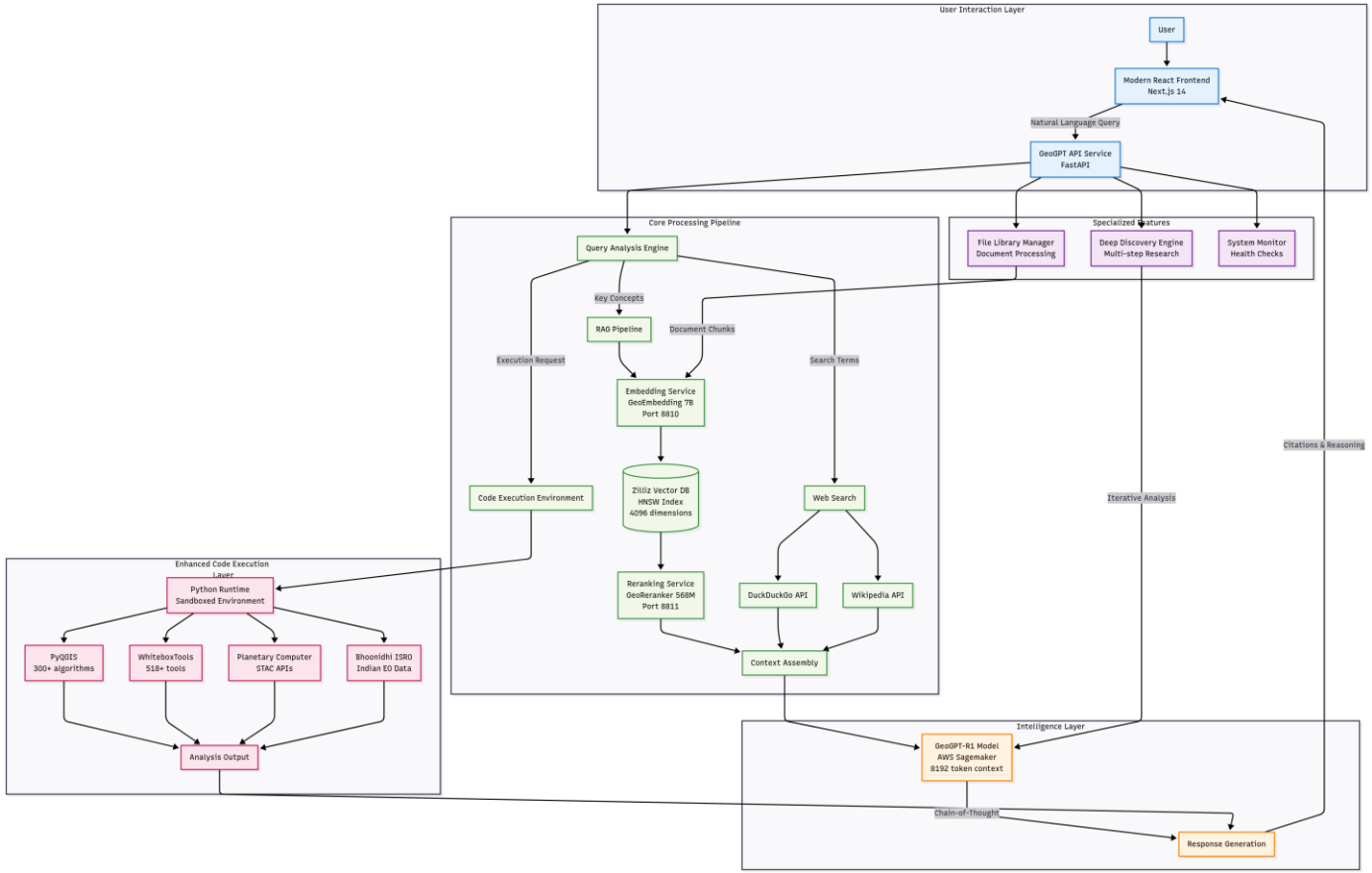
• **AI-written Discovery Report:** When the run completes, GeoGPT stitches findings, charts, and citations into a narrative you can export or continue chatting over.



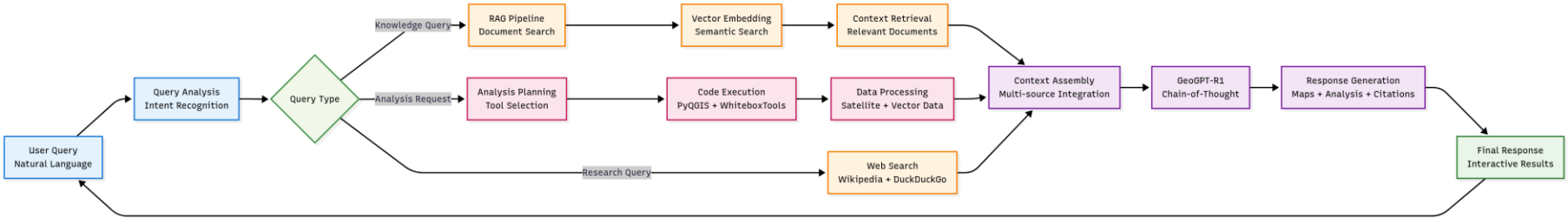
Feature: Deep Discovery Engine(UI)



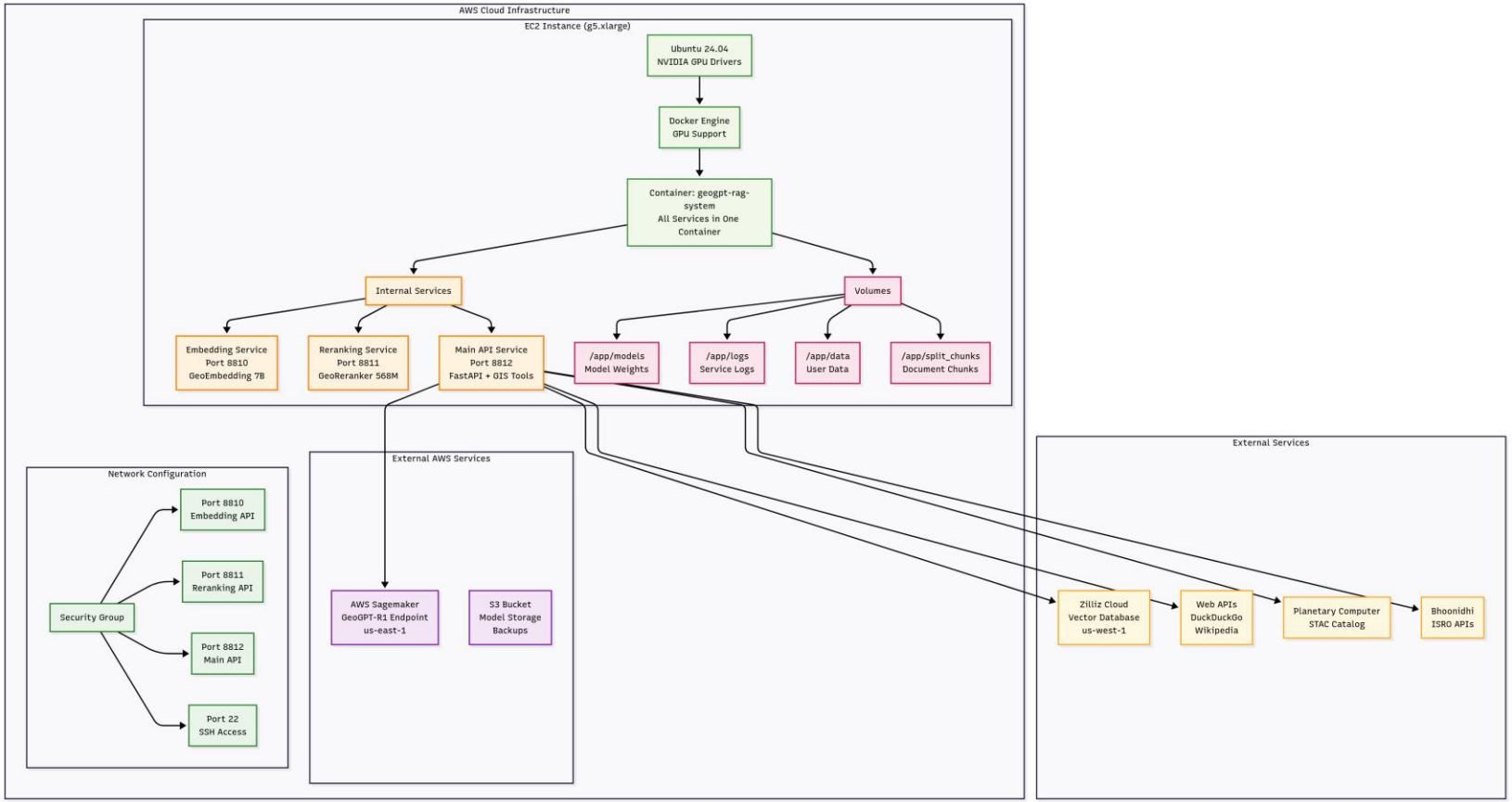
Architecture Diagram (Please Zoom IN)



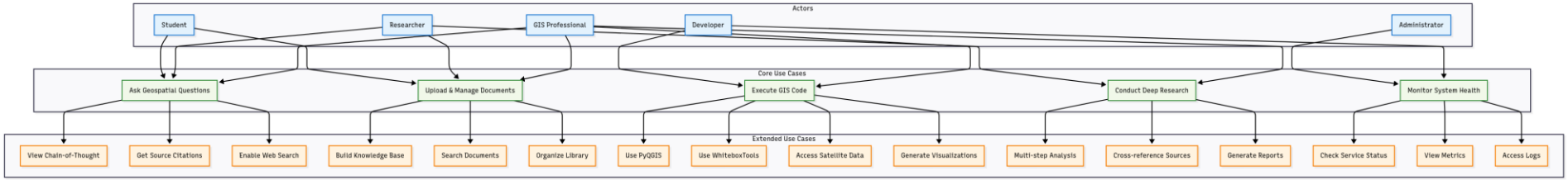
Process Flow Diagram (Please Zoom IN)



AWS Deployment Architecture Diagram (Please Zoom IN)



Use Case Diagram (Please Zoom IN)



Technologies Used (Technical README)

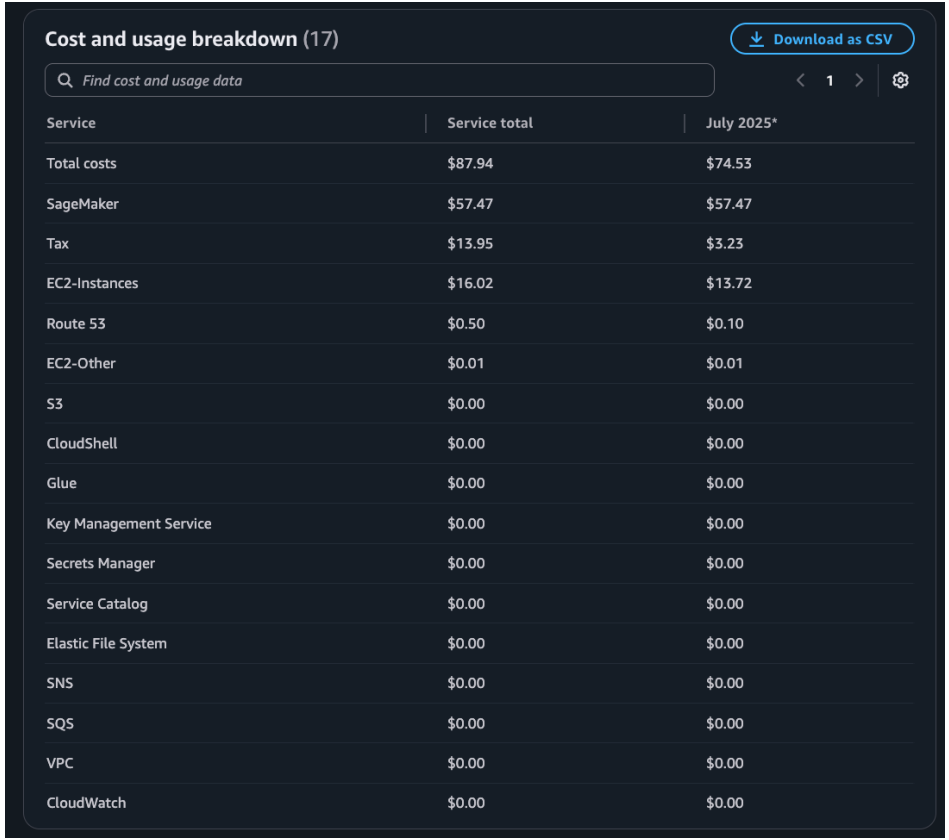
Layer	Key Tech	Why it matters
Frontend UI	**Next.js 14** · **React 19** · **TypeScript** · **Tailwind CSS 4** · **shadcn/ui**	Gives users a lightning-fast chat console and interactive map panels that work on any device without installs.
LLM & Reasoning	**GeoGPT-R1** (fine-tuned Qwen-1.5-7B) · **GeoReranker-568 M**	The brain of the app, it understands spatial language and picks the best answer steps with expert accuracy.
Embeddings & Vector Store	**GeoEmbedding-7B** · **Zilliz Cloud / Milvus v2.4**	Finds the right page or paragraph in millions of documents in under a fifth of a second.
Geoprocessing Engine	**PyQGIS 3.34** · **WhiteboxTools 2.3** · **GDAL 3.8**	Runs 800+ professional GIS algorithms. No expensive desktop software needed.
Data Access	**Microsoft Planetary Computer STAC** · **ISRO Bhoonidhi STAC** · Copernicus DEM · CHIRPS rainfall	Pulls both global and Indian satellite imagery on-demand so analyses are always up-to-date.
API & Micro-services	**FastAPI** · **gRPC** · **LangChain**	Streams answers in real time and lets internal services talk to each other.
DevOps & Infra	**Docker / Compose** · **AWS SageMaker** (LLM) · **AWS EC2 + ALB+VPC+EBS** (RAG) · **Terraform** · **GitHub Actions**	One-command deploys and blue-green updates with zero downtime keep the team shipping fast.

Implementation Cost

Cost Component	Unit Price (July 2025)	Hackathon PoC(≈ 1 month)	Production (per month)
LLM Inference GPU (SageMaker g6e.12xlarge)	~$4.34 / hr	**10 hr testing → $43.40**	720 hr (24×7) → $3 124.8
RAG / API Node (EC2 g5.xlarge)	~$0.97 / hr	**1 hr / day × 30 d → $29.10**	24×7 uptime → $698
Vector Store (Zilliz Cloud)	Free tier 1 M vectors · Standard $99 / mo	Free tier → $0	Standard tier → $99
Object Storage (Amazon S3 Standard)	$0.023 / GB / mo	**10 GB → $0.23**	1 TB → $23
Satellite Imagery Access	Planetary Computer = $0 Bhoonidhi ≈ $2 / scene	**≈ 12 scenes → $24**	~50 scenes → $100
Monthly Total	–	**≈ $97**	**≈ $4 050**

Costs can further be cut by utilizing smaller computer for endpoint inference, essentially deploying a smaller model.
The project is currently compatible with **Llama3.1-70B, Qwen2.5-72B** and **Deepseek-R1-Preview**. All these aforementioned models are trained on geoscience-specific subset of CommonCrawl.

Implementation Cost



Cost and usage breakdown (17)

Download as CSV

Find cost and usage data

1

Service	Service total	July 2025*
Total costs	$87.94	$74.53
SageMaker	$57.47	$57.47
Tax	$13.95	$3.23
EC2-Instances	$16.02	$13.72
Route 53	$0.50	$0.10
EC2-Other	$0.01	$0.01
S3	$0.00	$0.00
CloudShell	$0.00	$0.00
Glue	$0.00	$0.00
Key Management Service	$0.00	$0.00
Secrets Manager	$0.00	$0.00
Service Catalog	$0.00	$0.00
Elastic File System	$0.00	$0.00
SNS	$0.00	$0.00
SQS	$0.00	$0.00
VPC	$0.00	$0.00
CloudWatch	$0.00	$0.00

Video Demonstration Link

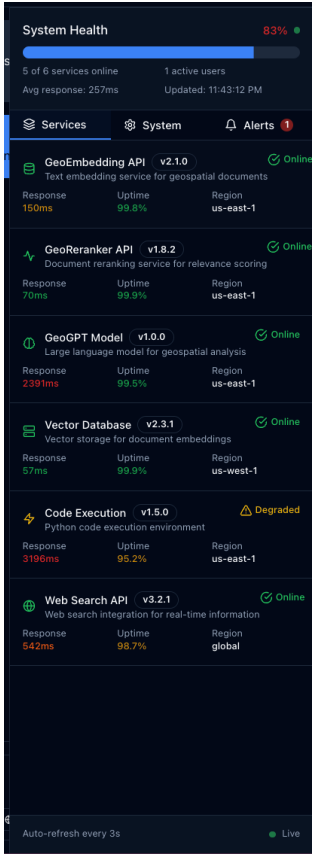
Youtube [https://youtu.be/6BRW9duBAfY]

Github Repository Link

Rekklessss/geogpt-rag
(https://github.com/Rekklessss/geogpt-rag)

Forgot to mention about system health monitoring and health service
Each port from the pipeline is added with a /health check to monitor the status of the APIs
Video demonstration doesn't include this too! (Completely Forgot)

LIVE MONITORING OF SERVICES



Mindmap (Bullets)

1. Selected an open source model (GeoGPT-R1-Preview) due to its high accuracy and COT process.
2. Deployed the model on an AWS Sagemaker Endpoint after creating a Huggingface TGI Container.
3. Hosted an EC2 Instance for RAG pipeline and Middleware CORS.\
4. Implemented basic API calls for /add_file and /query for basic RAG file processing, this is where I added the two models for AI Intelligence inside RAG (GeoEmbedding and GeoReranker).
5. Created a Next.js frontend with the planned features and chat interface.
6. Integrated the pipeline with another Docker container for Sandboxed code execution and tool flow (PyQGIS, WhiteboxTools)
7. Added STAC APIs and Bhoonidhi Satellite imagery data in the pipeline processing.
8. Pipeline monitoring, deployment scripts, documentation

Geo

Powered by H2S

BHARATIYA ANTARIKSH HACKATHON 2025

THANK YOU