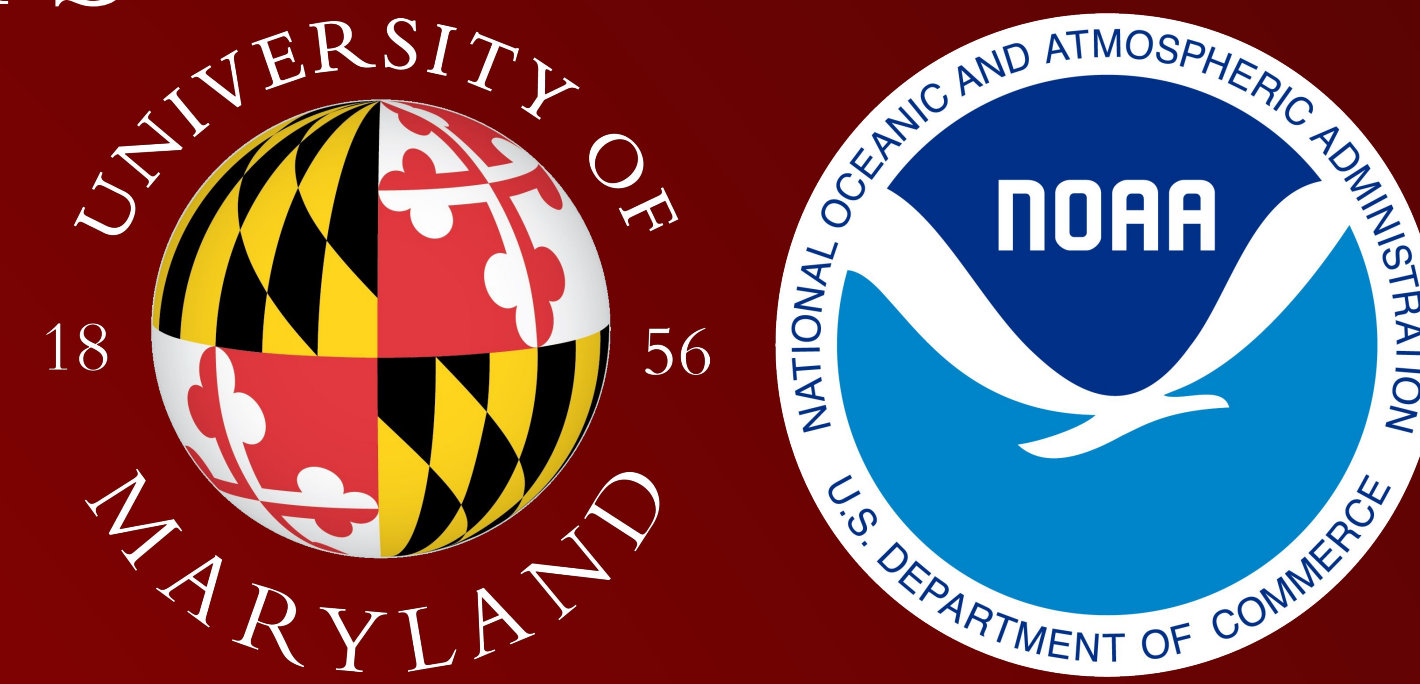


An Agentic LLM Framework for Fast and Interpretable Data Queries: Application to STOFS

Aryan Harooni¹, Hooman Moghimi², Atieh Alipour³ and Saeed Moghimi³.

¹ University of Maryland-College Park, ² Robert Frost Middle School, ³ Office of Coast Survey, NOAA



Abstract

We present a three-agent framework that combines function-calling large language models (LLMs) with high-level orchestration for fast and interpretable data access. The **Orchestrator Agent** manages dialogue and delegates tasks, the **Retrieval Agent** executes precise queries on structured datasets, and the **Analyst Agent** performs visualization and statistical evaluation. Applied to the Surge and Tide Operational Forecast System (STOFS), this architecture enables even non-experts to explore forecasts through natural language, reducing technical barriers and accelerating analysis.

System Architecture

Our framework combines three specialized large language models (LLMs) to enable agentic data querying:

Separation of concerns:

- **Orchestrator Agent** manages goals, interprets results, and engages in dialogue.
- **Retrieval Agent** executes structured queries via callable functions.
- **Analyst Agent** transforms outputs into plots, metrics, and statistical insights.

Workflow:

1. User issues a natural-language query.
2. Orchestrator interprets intent and delegates tasks.
3. Retrieval Agent accesses datasets (raw/post-processes) with precise function calls.
4. Analyst Agent evaluates outputs and generates interpretable visuals/statistics.
5. Orchestrator synthesizes results into a narrative for the user.

Advantages:

- Efficiency in accessing structured datasets.
- Interpretability via natural-language explanations and visual outputs.
- Human-in-loop flexibility to refine or steer results.
- Accessibility for non-expert users.

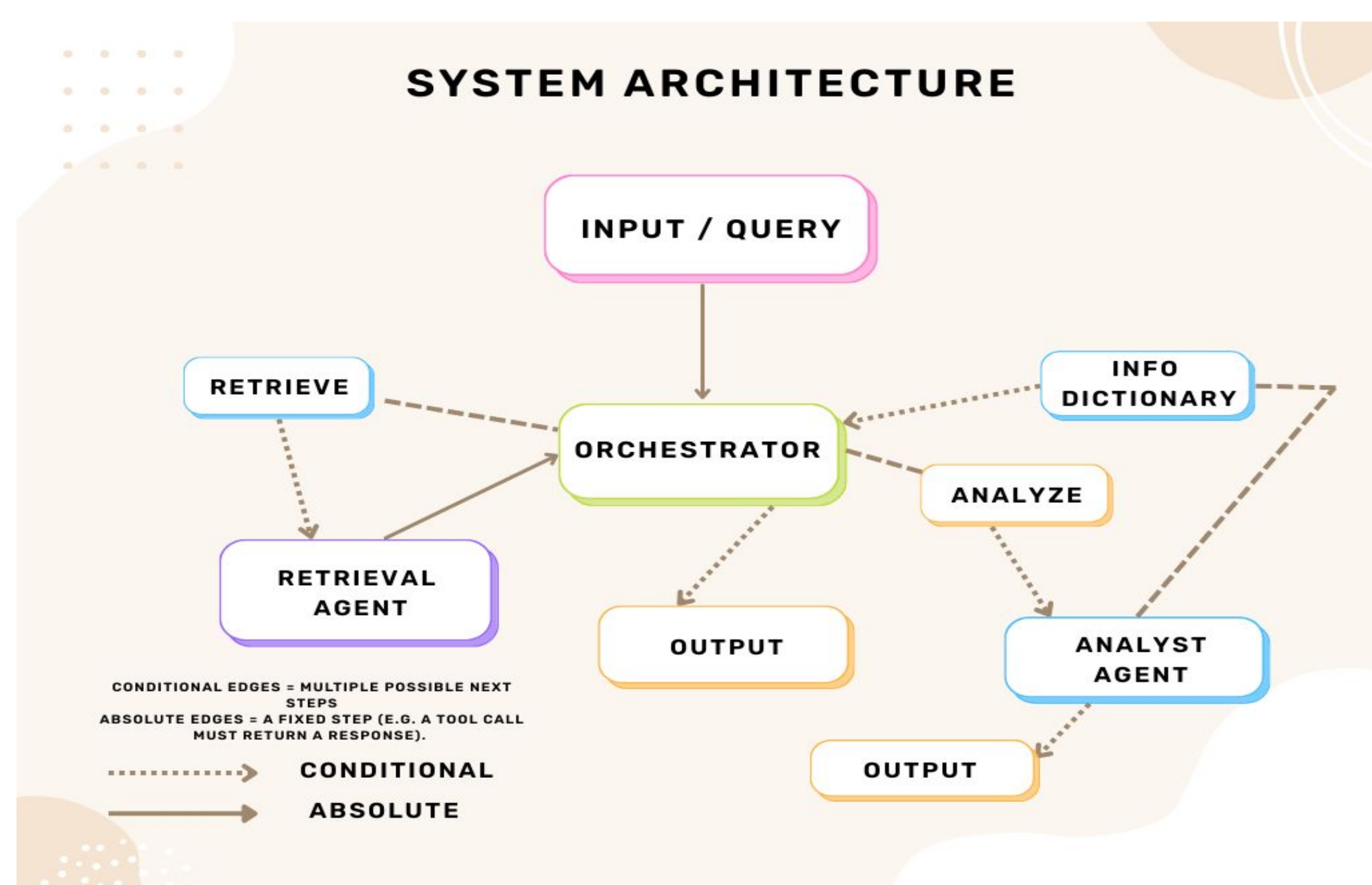


Figure 1. Three-agent system: Orchestrator manages queries, Retrieval fetches data, Analyst generates insights.

Orchestrator Agent

- ❖ **Role:** High-level coordinator that reasons over queries and results.
 - Parses user intent from natural-language.
 - Delegates tasks to Retrieval and Analyst agents.
 - Integrates outputs into coherent, human-readable responses.
 - Maintains conversational flow across multiple queries.
- **Strengths:**
 - **Abstraction** — shields users from technical details.
 - **Readability** — translates data into plain language insights.
 - **Adaptability** — manages multi-step and comparative workflows.
 - **Collaboration** — incorporates user feedback mid-process to refine answers.

Retrieval Agent

- ❖ **Role:** Interfaces with structured data through callable functions.
 - Executes parameterized queries (date ranges, locations, stations, metrics, etc.)
 - Handles data preprocessing (filtering, aggregation).
 - Ensures accurate, reproducible access to water-level outputs.
- **Strengths:**
 - **Precision** — avoids ambiguity by mapping natural language to exact queries.
 - **Modularity** — functions extendable to new datasets.
 - **Transparency** — every data call is logged and inspectable.

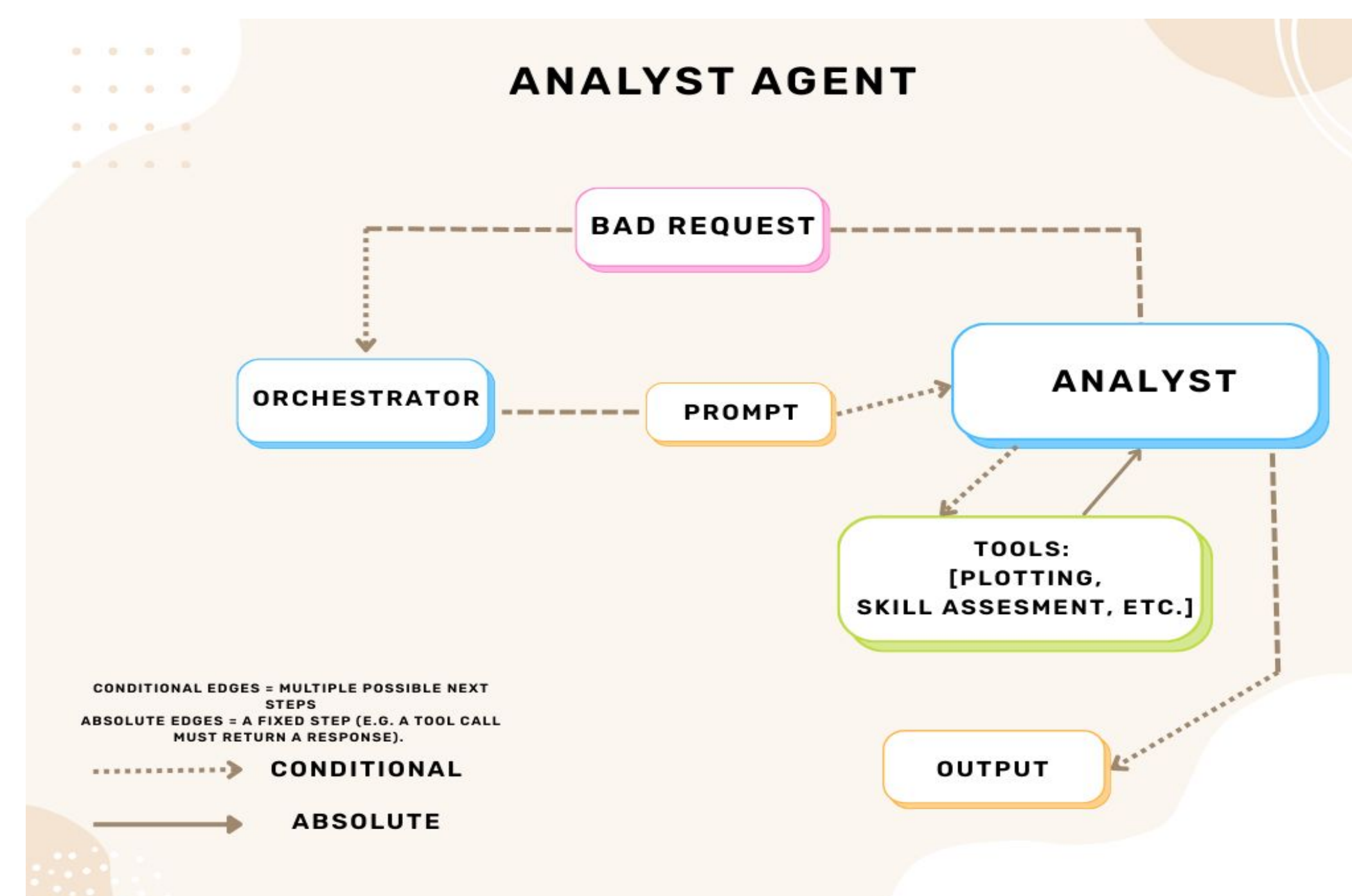


Figure 2. Retrieval Agent: Fetches raw/processed data, and flags invalid requests.

Analyst Agent

- ❖ **Role:** Converts retrieved data into interpretable insights.
 - Generates plots (time series, error dist., skill scores)
 - Performs statistical analyses (bias, RMSE, correlations, etc.)
 - Assesses model performance and highlights patterns.
- **Strengths:**
 - **Interpretability** — bridges raw numbers and useful insights.
 - **Flexibility** — supports both exploratory and standardized analysis.
 - **Visualization** — simplifies understanding through graphical outputs.

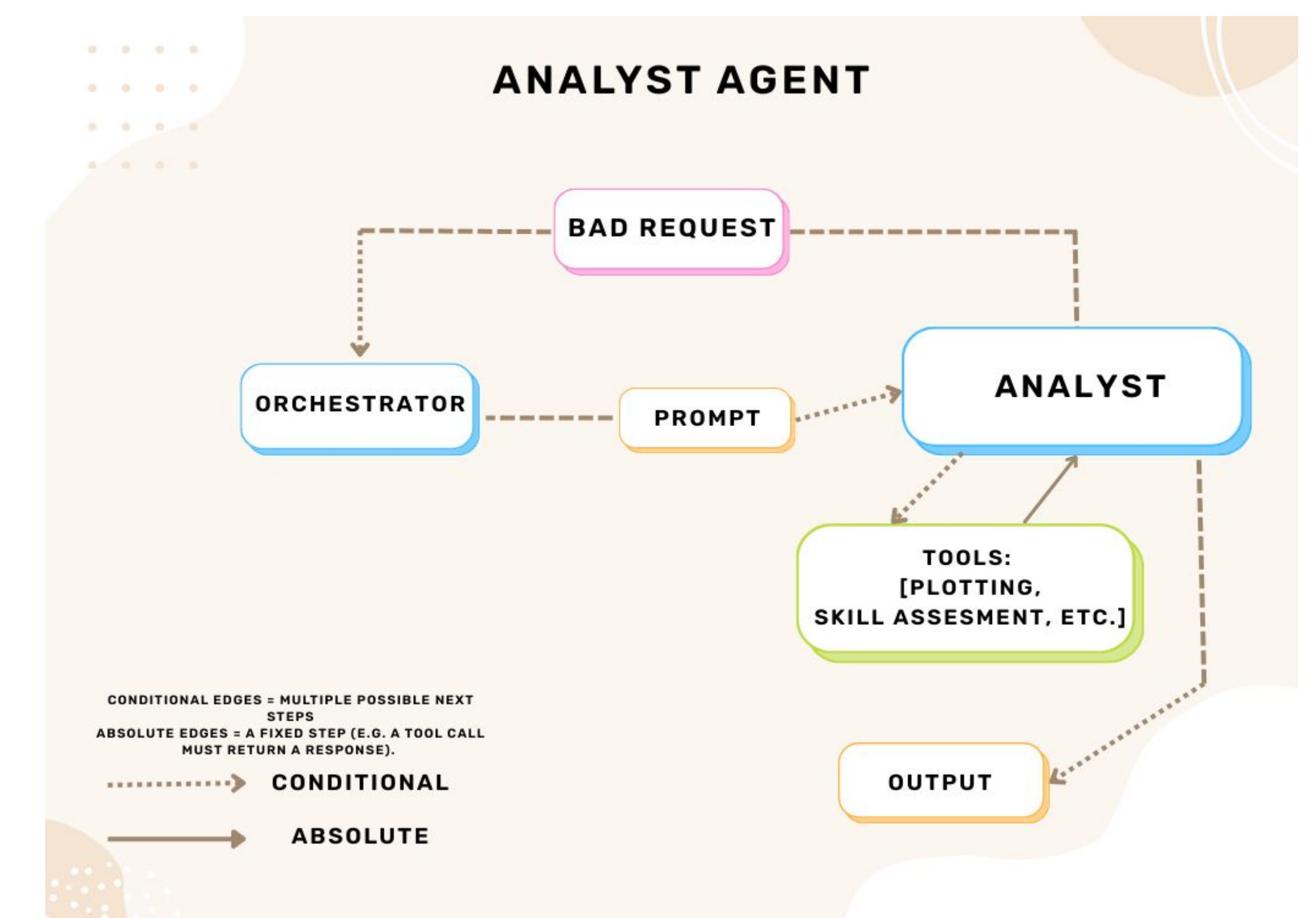


Figure 3. Analyst Agent: Generates plots, statistical evaluation, and skill assessment from the retrieved raw data.

Sample Chat

```
User: hey what can you do for me?
System: I can answer questions, retrieve STOFS data, run analyses, and chat.

User: tell me max RMSD on Jan 20, 2025
System: Maximum RMSD = 41,417.81 (Station 8571421, Lead = 144 hrs)

User: what even is RMSD?
System: RMSD (Root Mean Square Deviation) measures how much values deviate from a reference. Often used in molecular dynamics.

User: what does RMSD mean in water levels?
System: In this context, RMSD quantifies the difference between predicted and observed water levels. Lower RMSD = more accurate forecasts.

User: what are some other statistical analyses?
System: MAE, MAPE, R², RMSE, AIC, and BIC, among others.

User: what have we talked about until now?
System: [Summarizes prior conversation]

User: quit
System: Bye.
```

Figure 4. User query answered with retrieval, explanation; plotting shown here is a feature under development.