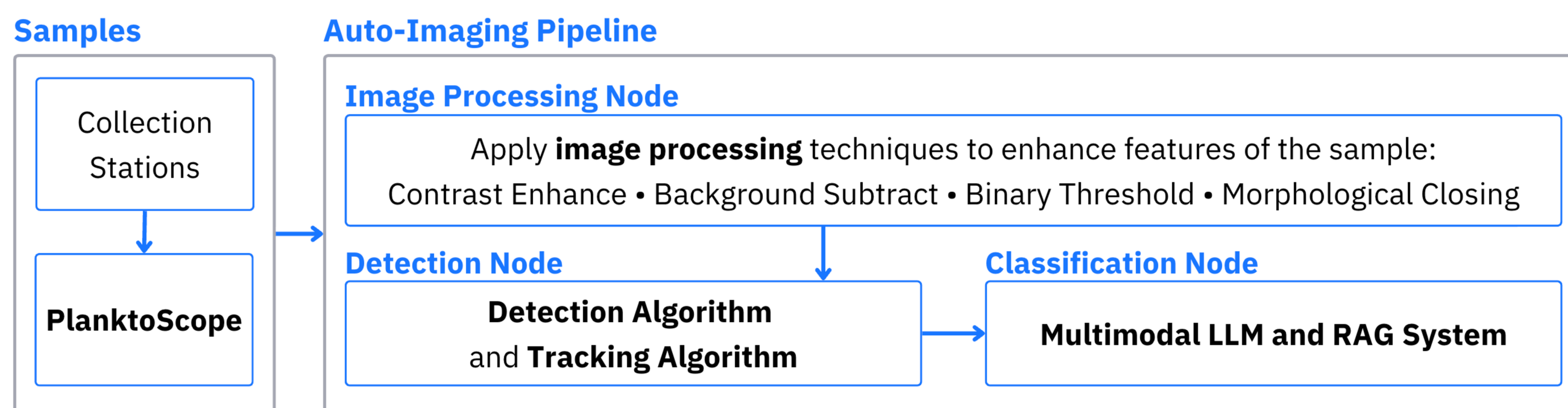
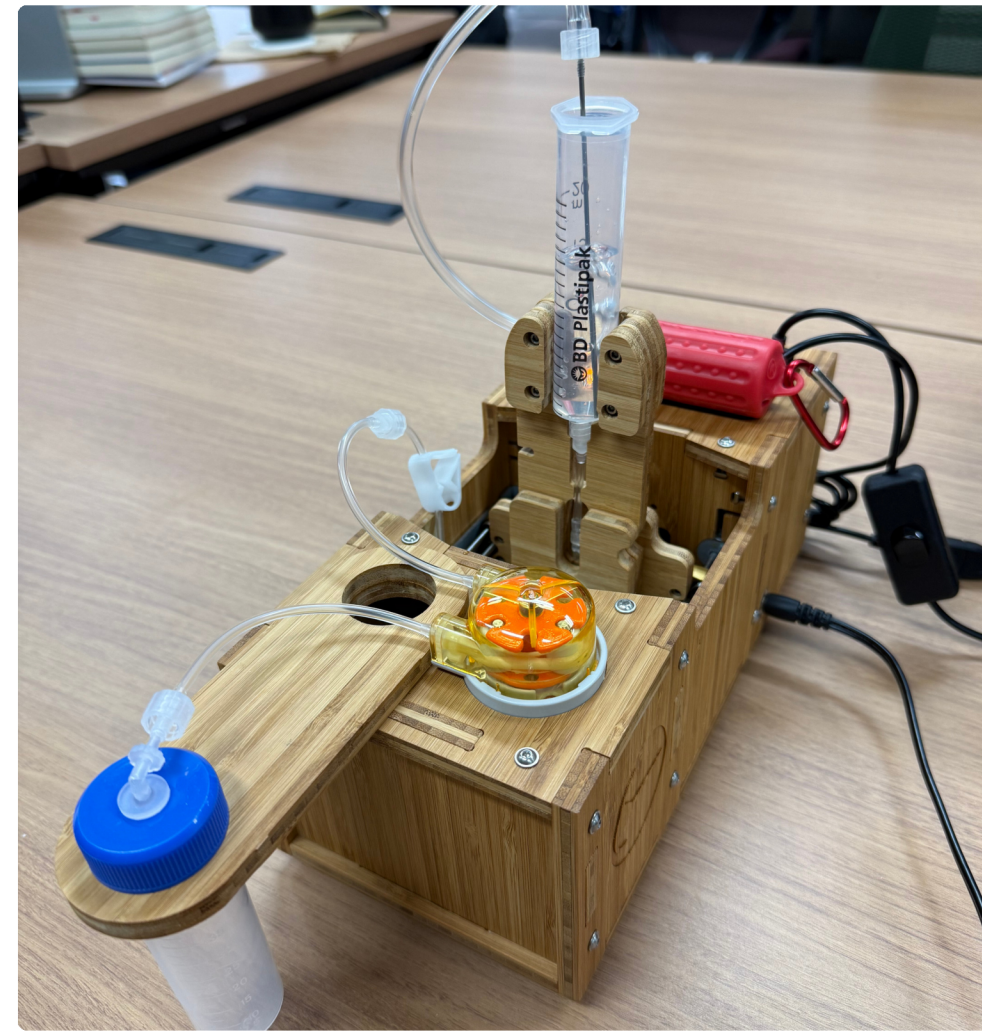


Context-Aware Marine Plankton Classification with Multimodal Large Language Model and Retrieval-Augmented Generation Reasoning

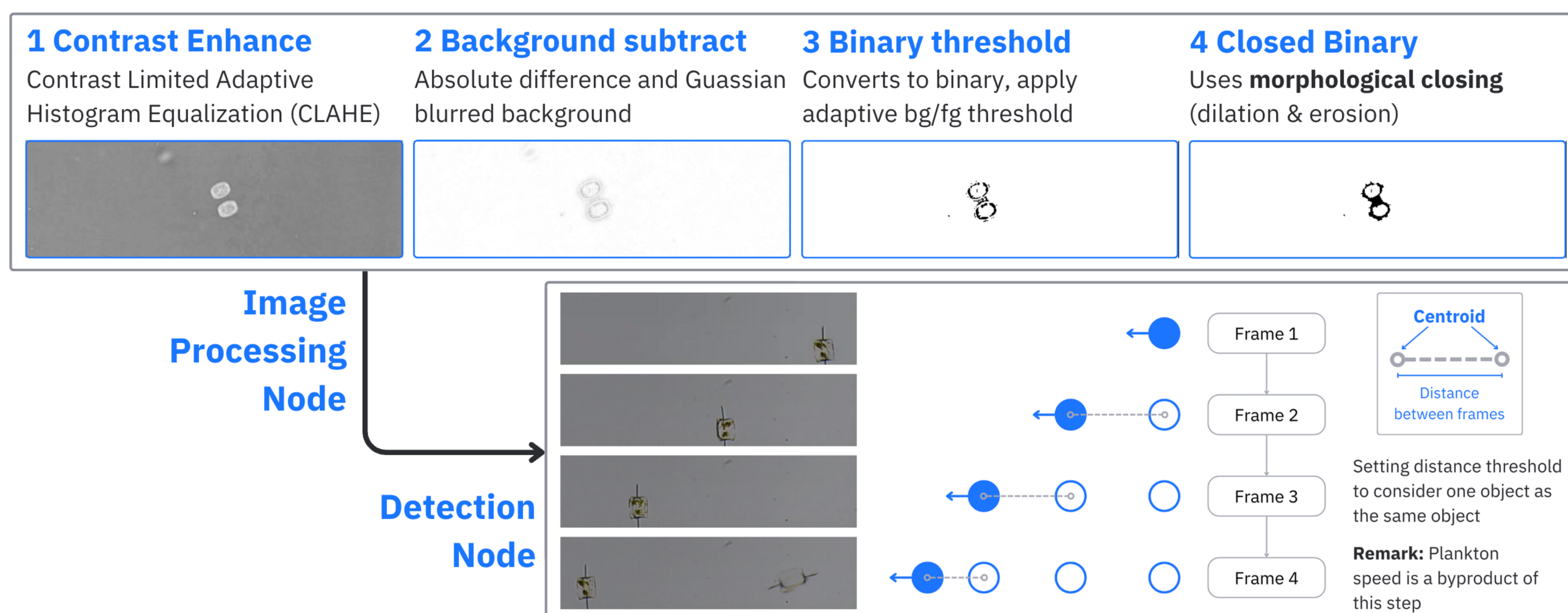
Jaronchai Dilokkalayakul¹, Akane Kitamura², Takeshi Obayashi^{1,2} (¹Graduate School of Information Sciences, Tohoku University | ²WPI-AIMEC)

Summary & Method Pipeline

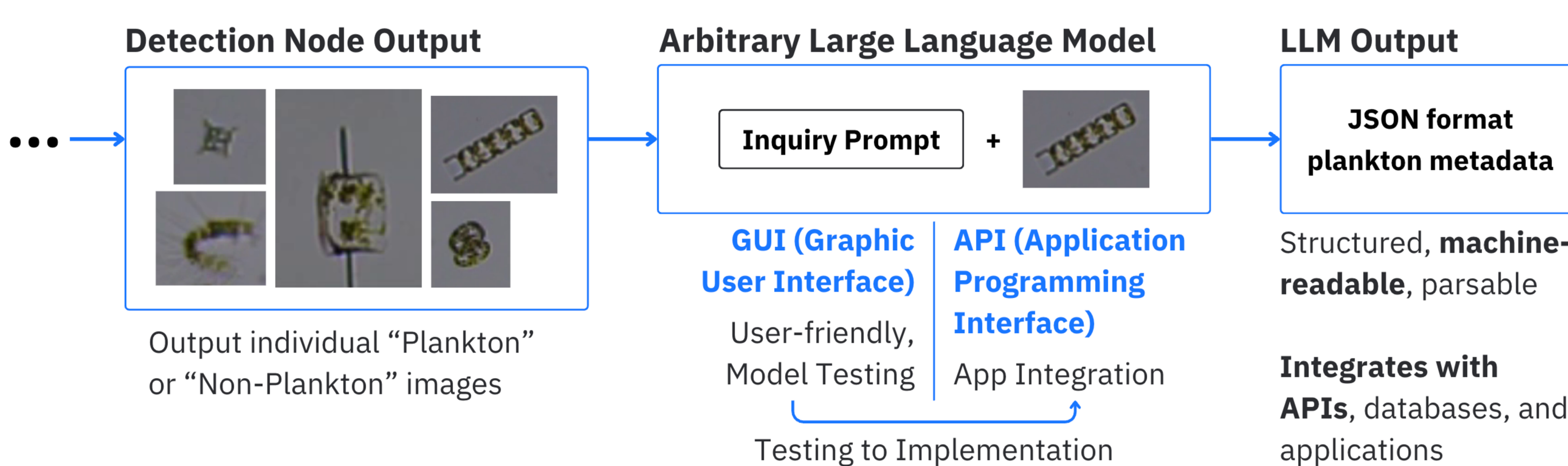
Traditional plankton identification methods are labor-intensive, posing challenges for monitoring marine ecosystems and environmental changes. While CNNs automate classification processes, CNNs depend on large labeled datasets. To address this, we propose a framework that combines large language models (LLMs) with retrieval-augmented generation (RAG) to automate plankton classification, by retrieving semantically similar examples from our curated vector database and integrating them into the LLM's input. This method is **context-aware**, **scalable classification with reduced reliance on labeled data**.



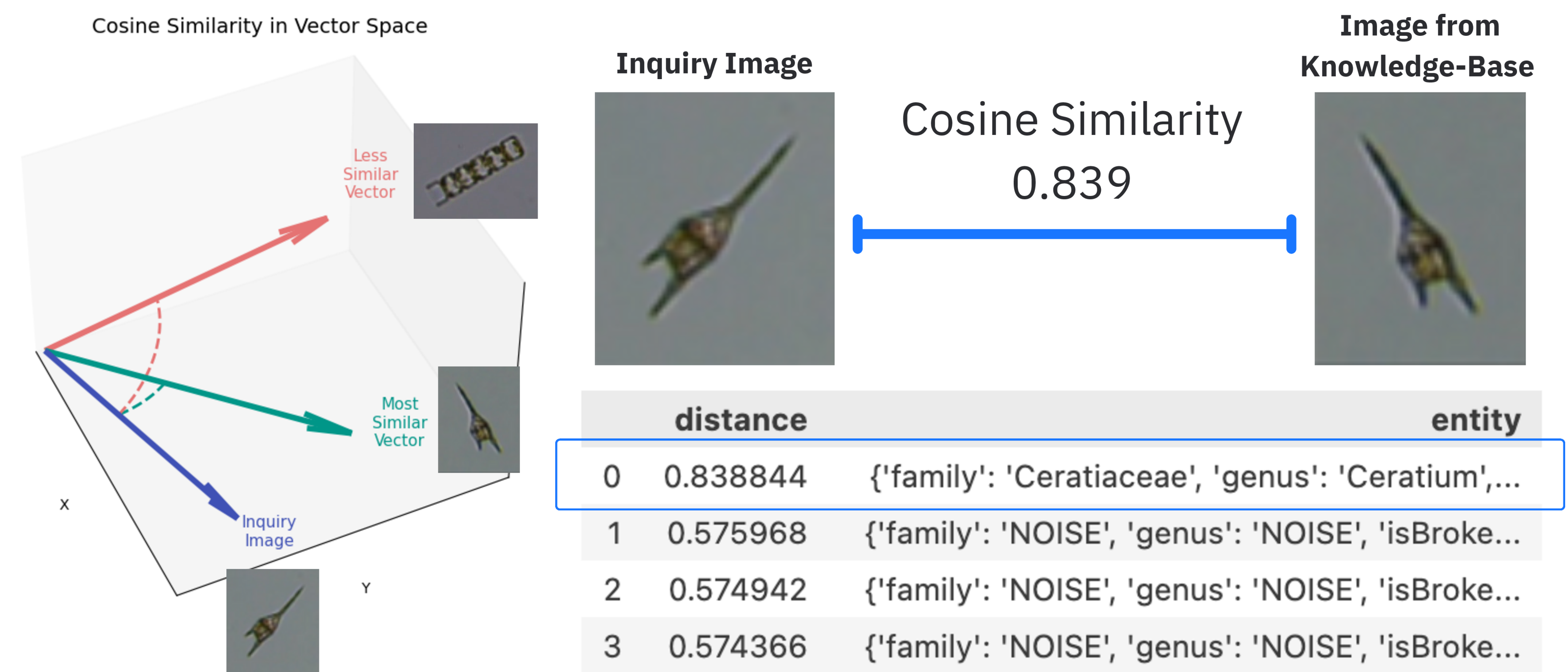
Auto-Imaging Pipeline



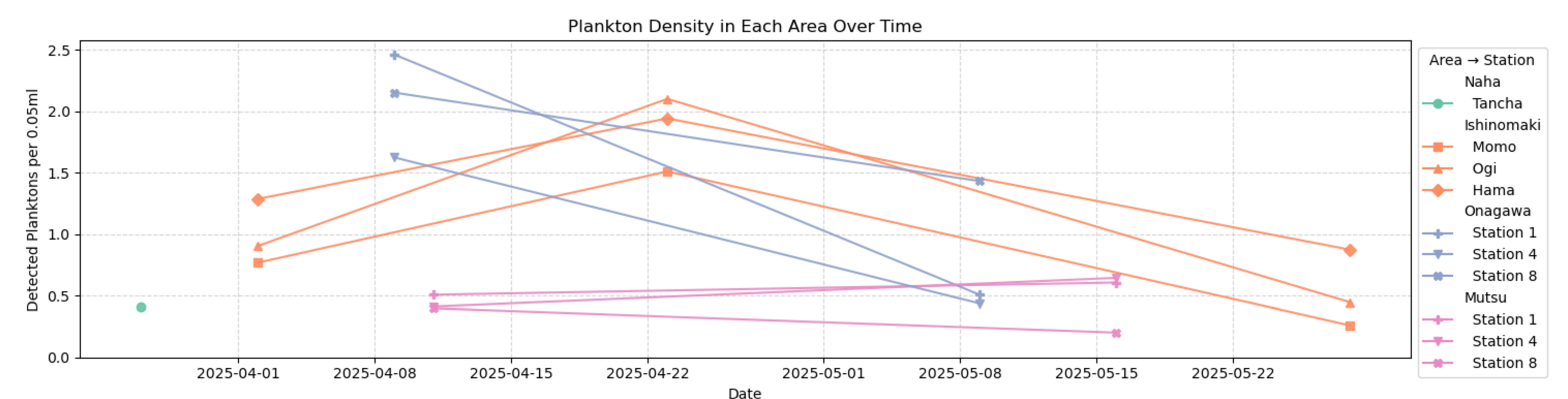
Key Concept of LLM Inquiry



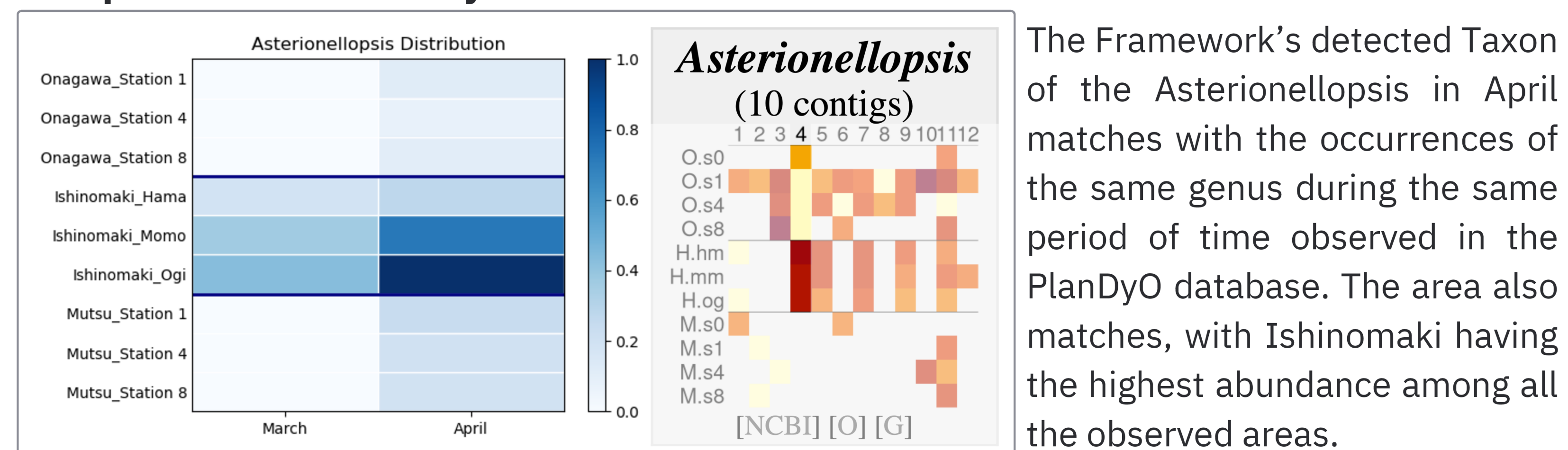
Curated Plankton Vector Database



Framework Monitoring Results



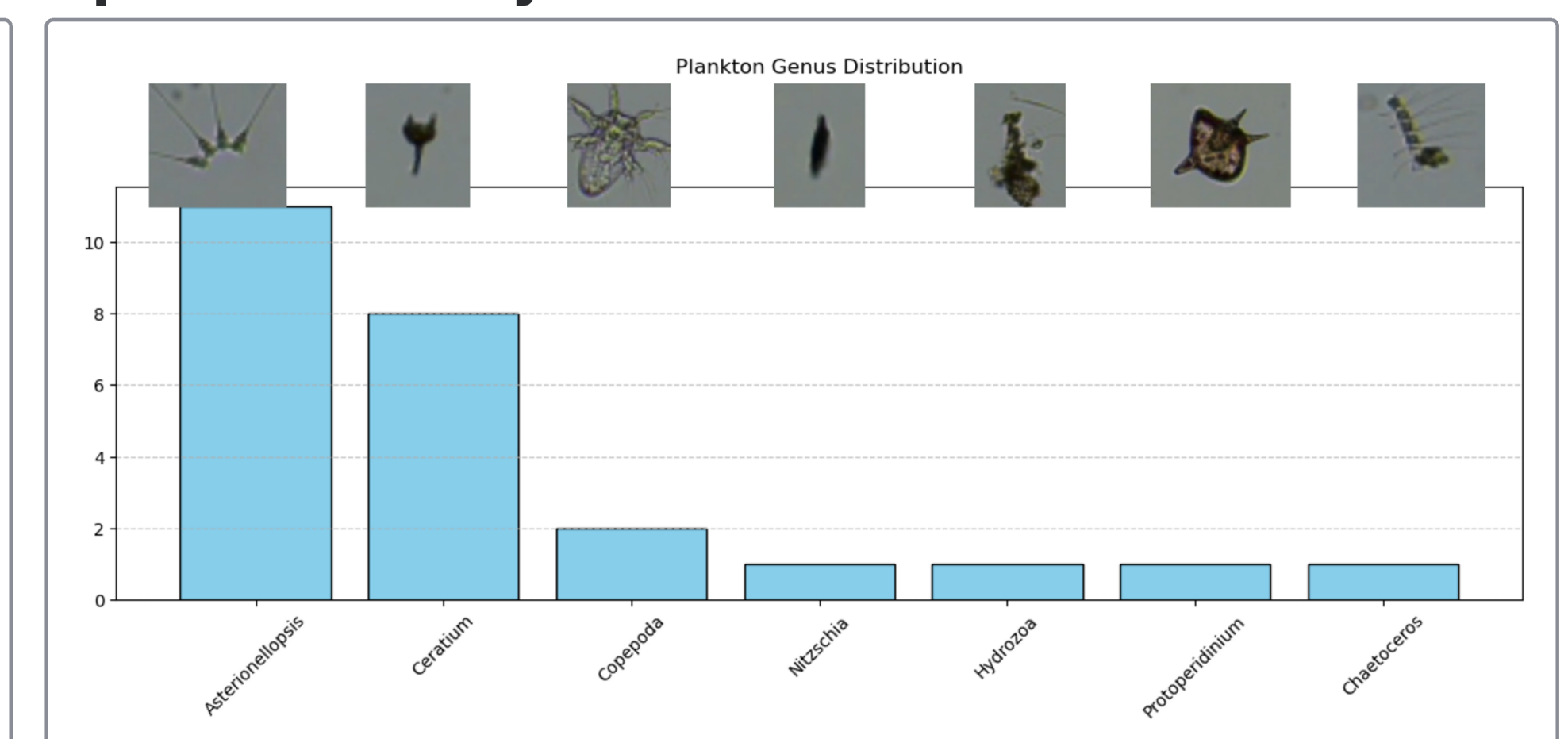
Comparison with PlanDyO Occurrence Data



Taxonomy Output

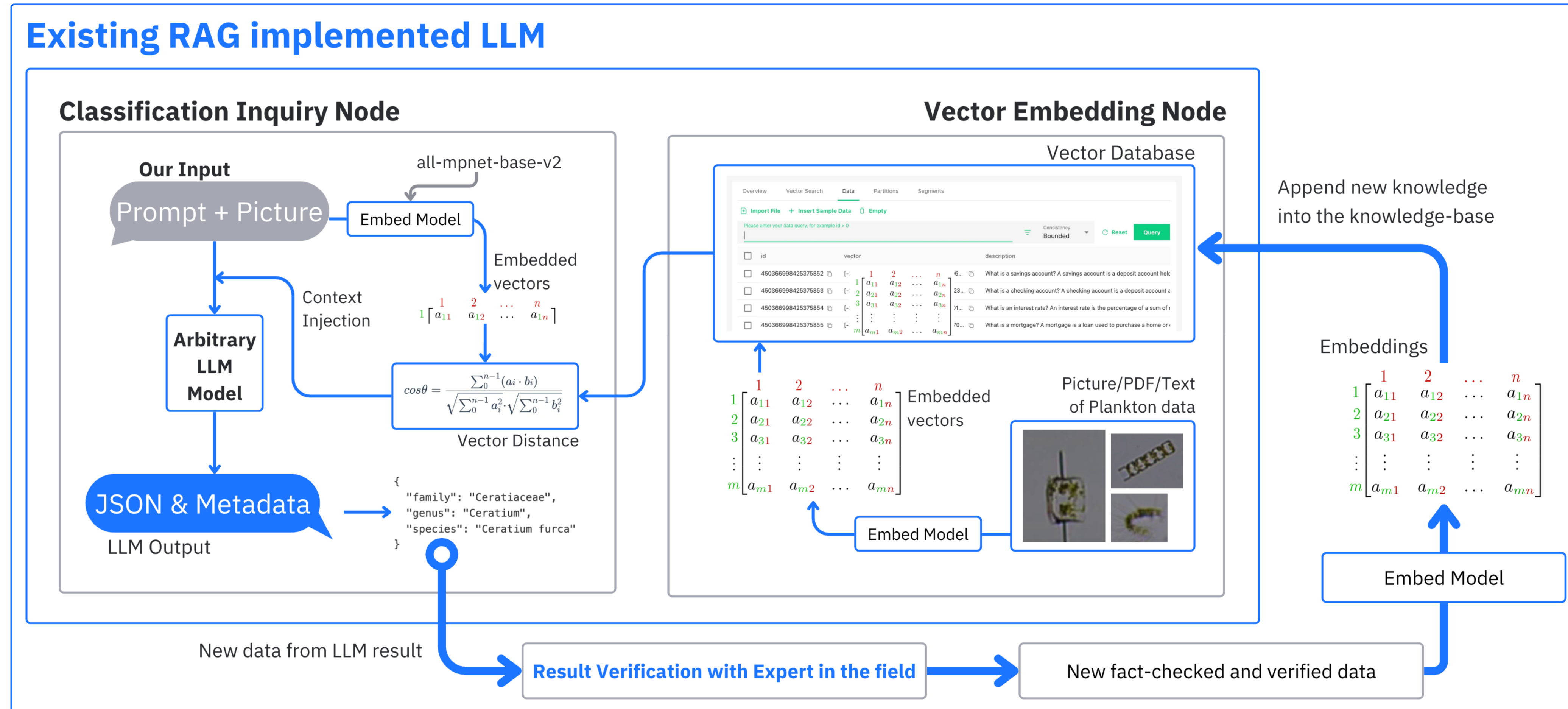


April Mutsu Bay Distribution



Large Language Model and Retrieval-Augmented Generation Framework

RAG implemented LLM with Feedback Loop



A prompt and an image of a plankton specimen are provided as the input query. Then the image and textual context are encoded into vector embeddings. The embedding is used to search a vector database containing embedded plankton documents, this retrieves semantically similar examples. The retrieved context is then injected into an LLM, which uses both the input and retrieved information to generate a structured JSON taxonomic metadata output. New outputs are verified by experts to confirm the result. Validated examples are embedded and added back to the vector database, as a continuous improvement to the system's contextual grounding.

Contact Information

Presenter: **Jaronchai Dilokkalayakul**

Email: dilokkalayakul.jaronchai.p8@dc.tohoku.ac.jp

Personal Website: <https://jaronchai.com>