5703 Group Based Capstone Project

Week 6 Tutorial Presentation

Project Number: CS55-1

Project Title: Agentic Multimodal RAG_An Intelligent Framework for Scientific Concept Discovery from Text and Visuals

Presenters: Hanyu Wang, Xiaoran Wang, Zhencheng Huang, Jinlin Zhong, Kunming Lyu, Junbo Liu

THE UNIVERSITY OF SYDNEY

Summary of the most recent client meeting

Xiaoran Wang (xwan5753)

- Date of Meeting: 03/09/2025
- Key Points Discussed:
  - Presented our incremental implementation plan for the multimodal DocQA project (text + figures/images + tables).
  - Agreed to keep evaluation objective using public QA datasets with ground truth (e.g. ScienceQA/DocVQA) and accuracy/quote-based metrics.
  - Baseline pipeline to demo: CLIP + text embeddings → multimodal retriever → open-source LLM → answer with citations.
  - Logistics: Slack remains primary comms; short demo/screenshots encouraged; GPU access can be requested via university.
- Client Feedback and Requests:
  - Keep increments small, clear, and demonstrable; focus on clarity and novelty.
  - Bring a concise PPT next meeting showing: problem statement, baseline pipeline, incremental steps & milestones/owners, dataset shortlist & evaluation plan, and risks.

Recap of Last Week's Plan

Kunmign Lyu, klyu0947

- Planned Tasks/Goals from Last Presentation:
  - Complete project proposal based on client feedback
  - Ensure alignment of deliverables with client's stated requirements
  - Prepare draft version for internal review before submission

Group Progress

THE UNIVERSITY OF SYDNEY

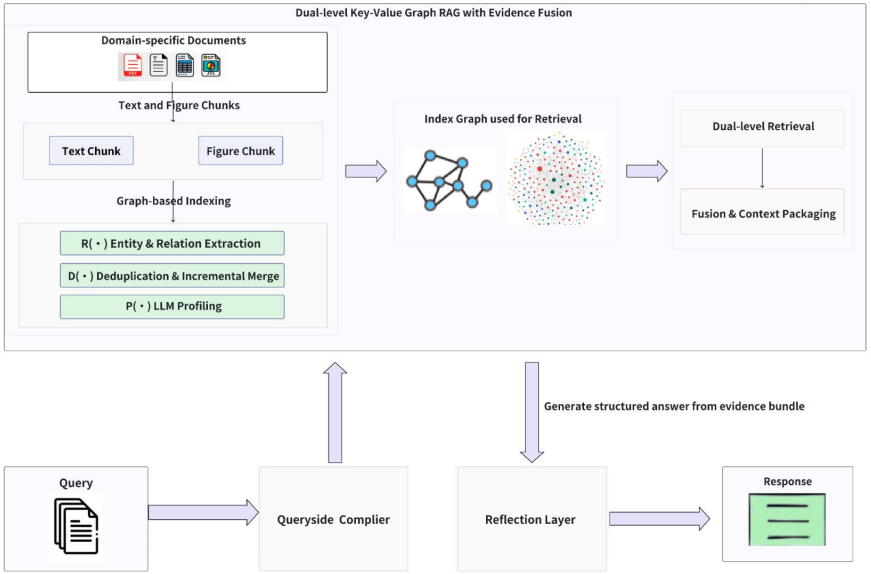
Kunming Lyu, klyu0947

- Achievements since Last Presentation
  - Completed project proposal
  - Refine the project workflow and finalize the techniques
  - Experimented with possible optimization points of LightRAG
  - Researched potential evaluation methods
- Alignment with Client/Project Requirements
  - Incremental design matches client expectations for clarity, novelty, and demonstrability.
  - Public QA datasets with ground truth ensure measurable and transparent evaluation
  - Initial demo provides a clear and testable starting point.

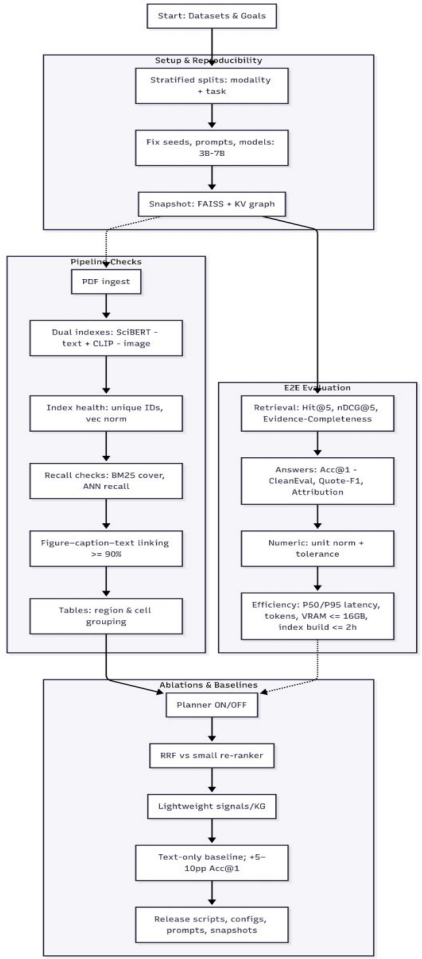
Group Progress Evidence

Kunming Lyu, klyu0947

Screenshot / Demo



to

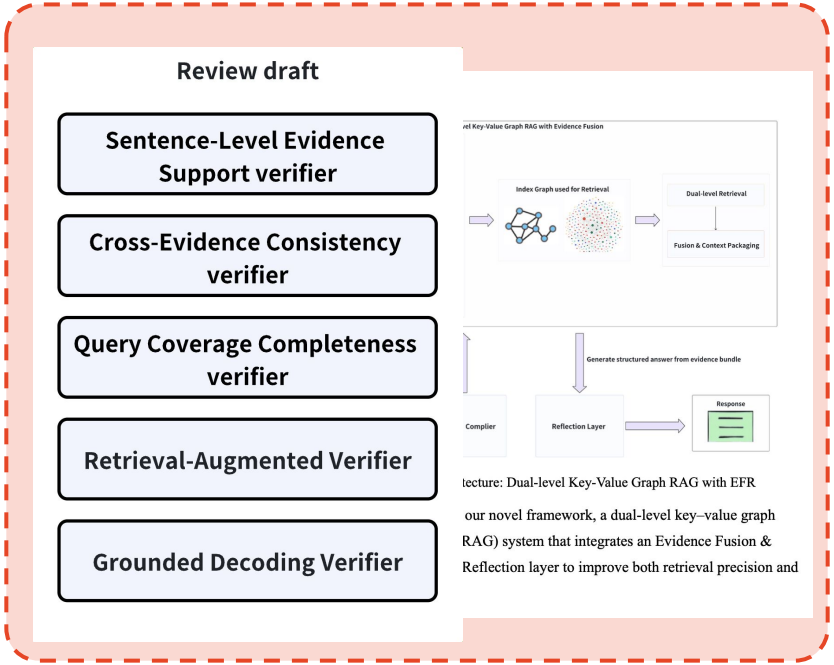


Individual Progress – Member 1

[Jinlin Zhong(Richard) jzho0102]

- Finish the Methods and Deployment section of the proposal
- Design and consolidation of the incremental Reflection architecture:
  - Sentence-Level Evidence Support Verifier
  - Cross-Evidence Consistency Verifier
  - Query Coverage Completeness Verifier
  - Retrieval-Augmented Verifier
  - Grounded Decoding Verifier

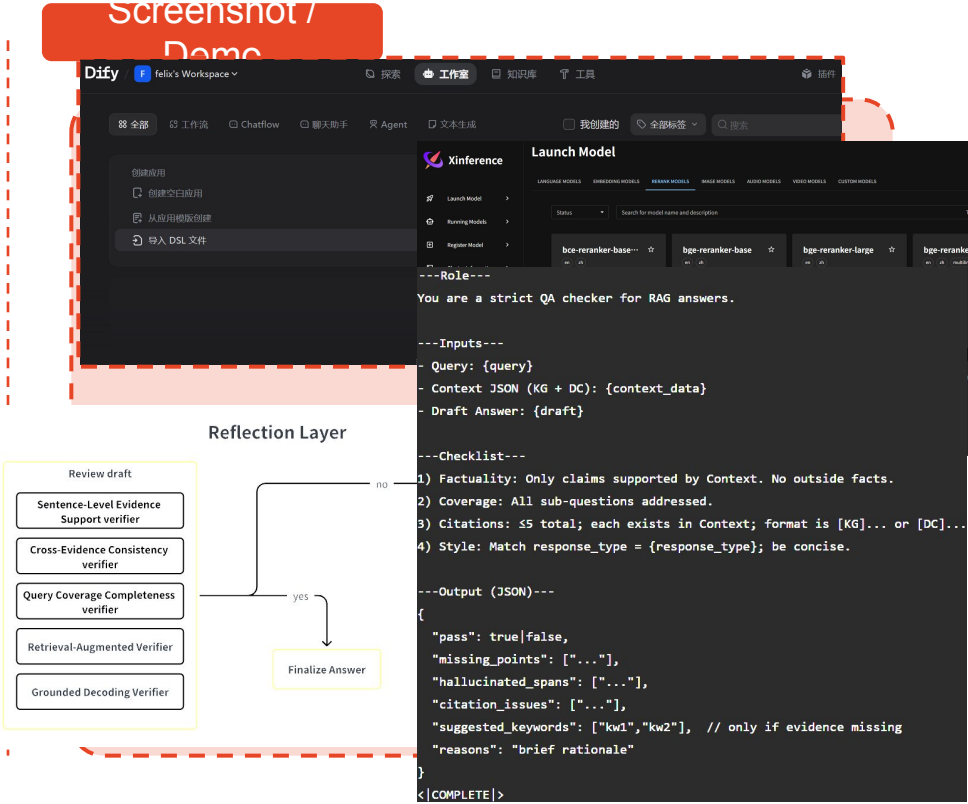
Screenshot / Demo



Individual Progress – Member 2

[Hanyu Wang, 540063972]

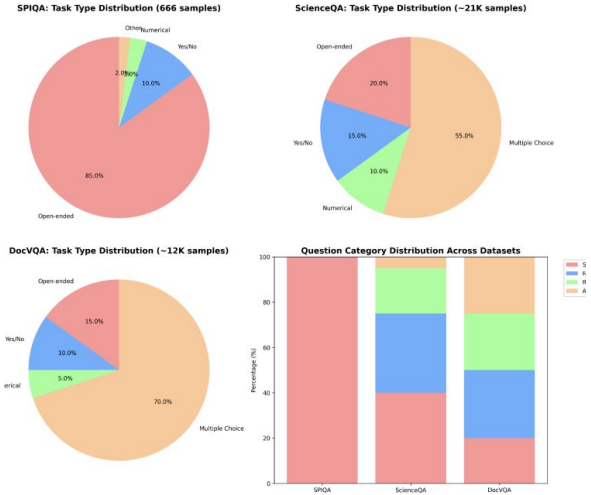
- Designed and completed my part of the system architecture
- Organized meetings and reviewed reports to drive project progress
- Successfully deployed Dify and Xinference in local environment
- Experimented with embedding the Reflection Layer into the existing framework
-

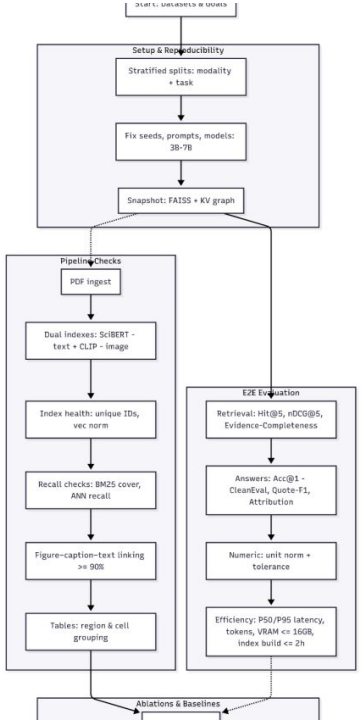


Individual Progress – Member 3

[Junbo Liu - jliu0232]

- Completed unified EDA on SPIQA, ScienceQA, and DocVQA
- Deployed and validated LightRAG end-to-end: PDF → indexing (text + image) → retrieval → answer; produced a working demo.
- Finished the Scope Proposal (data analysis & testing structure).
- Updated the repo with analysis notebooks, LightRAG configs, run scripts, and a README for reproducibility.



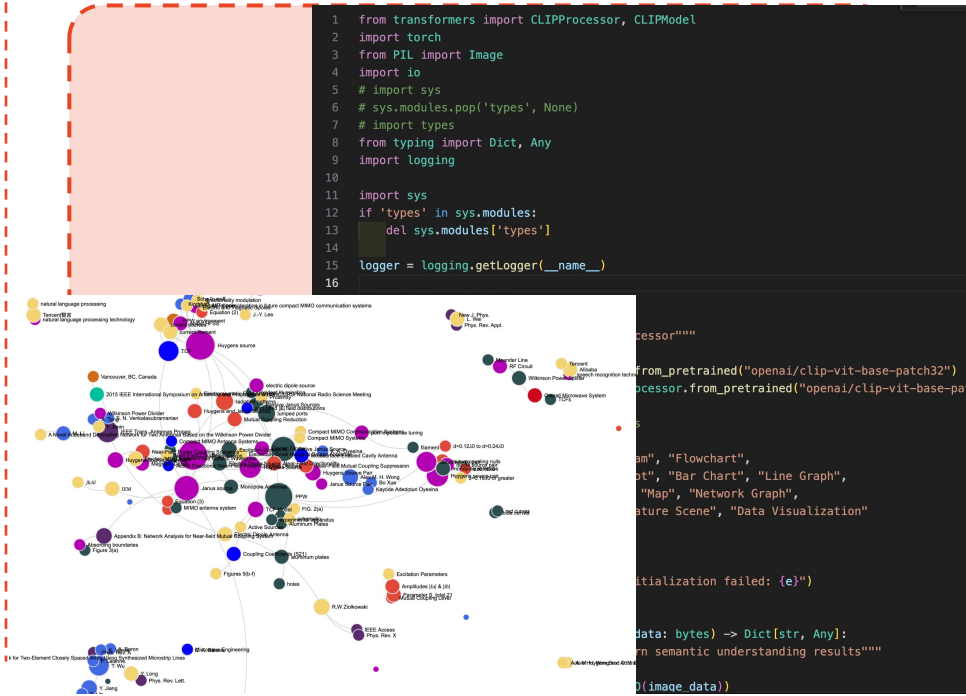


Individual Progress – Member 4

Kunming Lyu, klyu0947

- Completed part of the project proposal.
- Deployed LightRAG and gained understanding of its framework and workflow.
- Experimented with integrating CLIP into LightRAG to process images.
- Explored possible approaches for interpreting tables within LightRAG

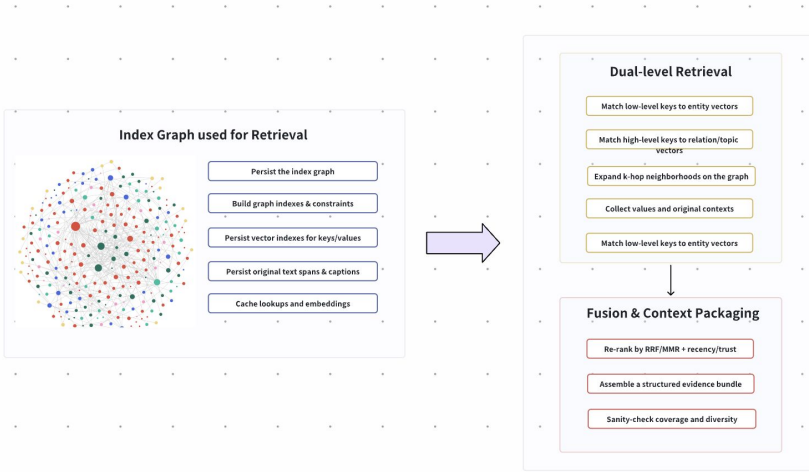
Screenshot / Demo

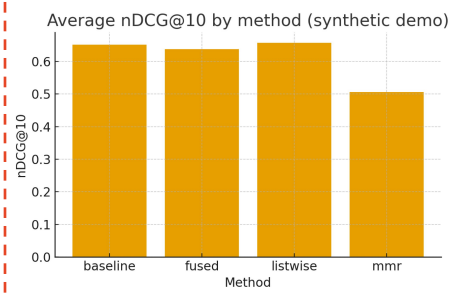


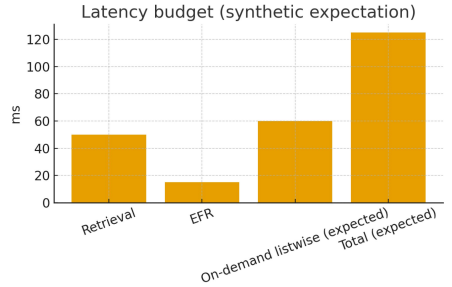
Individual Progress – Member 5

Zhencheng Huang, zhua0186

- Completed the EFR technical framework
  - Weighted RRF for multi-channel fusion (+ per-channel attribution).
  - On-demand listwise re-ranking with lightweight cross-encoder.
  - Post-re-rank MMR for de-dup & diversity.
- Demo findings
  - Pipeline runs end-to-end and is explainable.
  - Ablation (synthetic) — listwise stabilizes borderline cases.





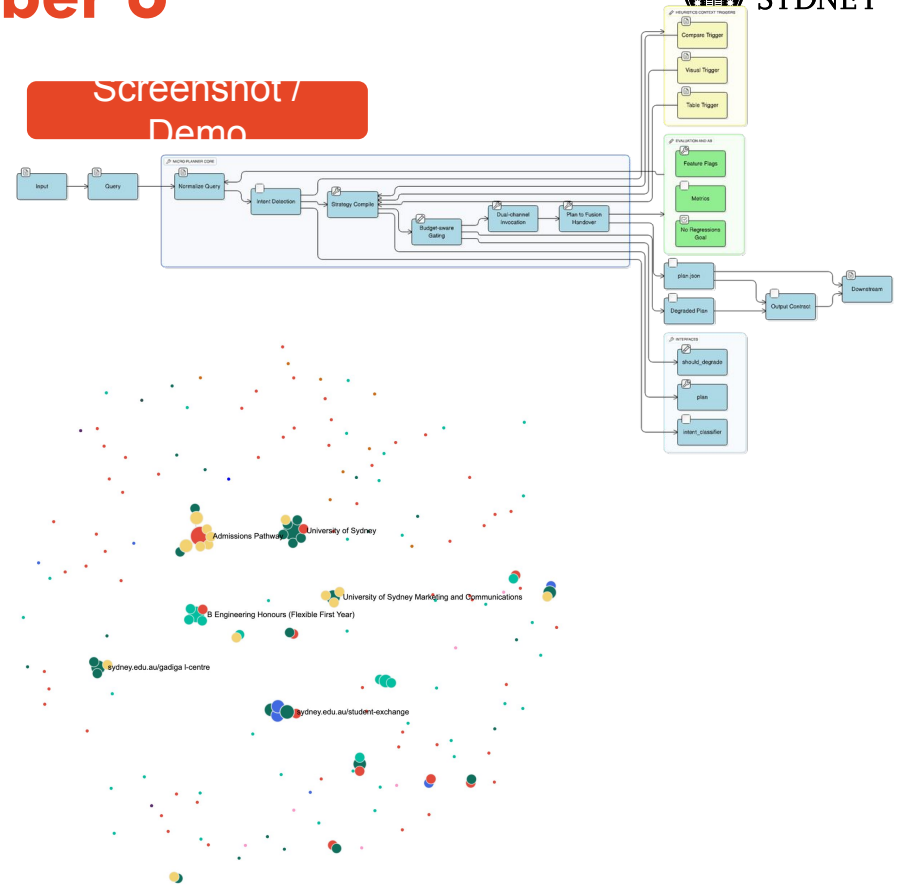


Individual Progress – Member 6

Xiaoran Wang (xwan5753)

- System Architecture Design
  - Designed a flow chart of our system architecture
  - Focused on integrating a Micro Planner to support adaptive tool routing in the pipeline.
- Implementation
  - Set up LightRAG demo environment on Mac with Ollama.
  - Configured .env, ingestion pipeline
- Documentation & Communication
  - Completed my part of the project proposal, Completed Client Meeting Minutes
  - Ready to prepare slides for next check-in.

Screenshot / Demo



Challenges & Issue

[Jinlin Zhong(Richard) jzho0102]

- Group-Level Challenges
  - Balancing accuracy and efficiency: Evidence fusion and re-ranking improve answer quality, but introduce additional latency and GPU/memory cost
  - Handling conflicting or incomplete evidence: Multiple sources may present contradictions, while some queries lack sufficient supporting data
  - Ensuring robust query understanding
  - Evaluation and verification: Automatic metrics (e.g., recall, NDCG) do not always align with human judgment of factuality and usefulness

What's the plan before next tutorial?

- Tasks/Goals for Upcoming Week
  - Significant progress of the increment architecture:
    - Evidence Fusion & Re-Ranking (EFR Layer)
    - Query-side Compiler
    - Micro Planner
    - Reflection Layer
    - Image and photo processing pipeline
  - Complete individual progress report


THE UNIVERSITY OF
SYDNEY