Faculty of Media Engineering and Technology
German University in Cairo
Dr. Nourhan Ehab

Presentation and Evaluation Guideline

Milestone 3 marks the final and most technically demanding stage of the project, where each team must demonstrate a fully integrated and functional Graph-RAG system. To ensure a fair and consistent assessment, the evaluation for Milestone 3 will follow a structured format and will be conducted **per team over a 45-minute slot**.

The evaluation consists of **two parts**:

1. **Team Presentation (Deliverables Evaluation)** ***15%***
   The team will present the implemented system, focusing strictly on the system architecture, pipeline integration, retrieval strategies, experimental results, and the live demo. This part evaluates the functionality, completeness, and performance of your Graph-RAG pipeline. The presentation should take a minimum of 18 minutes and a maximum of 22 minutes.

2. **Individual Q&A Evaluation** ***15%***
   Each team member will be examined individually on **their specific component**. Depending on the remaining time after the presentation, the Q&A may include **multiple rounds per person**. You may be asked to walk through portions of your code, explain the sequence of operations within your component, discuss error cases based on your experiments, or justify implementation decisions.
   When a question is directed to a specific member, **only that member is allowed to answer**.

This document outlines the required presentation structure, expectations for both parts of the evaluation, and the rules that will be applied. Please read it carefully and prepare your work accordingly to ensure a smooth and efficient evaluation process.

Faculty of Media Engineering and Technology
German University in Cairo
Dr. Nourhan Ehab

Presentation Outline:

- High-Level System Architecture (2 minutes)
  - Show an overview of your pipeline
  - Present your task of choice, and if you used an external dataset.
- Input Preprocessing (2 minutes)
  - **Intent classifier** (rule-based, LLM-based, or hybrid)
  - Entity extraction, with examples
  - Embedding step (if used)
- Graph Retrieval Layer - Baseline (2-3 minutes)
  - Show **your Cypher query templates (at least 10 should be implemented)**
  - Show snippet of retrieved nodes/relationships
- Graph Retrieval Layer - Embedding-Based Retrieval (2-3 minutes)
  - State the approach you selected: **Node embeddings** OR **Feature vector embeddings**
  - Show the two embedding models compared, with experiment results
- LLM Layer (3-4 minutes)
  - **Context Construction**, how you integrate the input, baseline Cypher query output, and the embedding output.
  - Prompt Structure
  - LLMs Comparison; experiments should include quantitative metrics and qualitative evaluation.
- Error Analysis & Improvements (2 minutes)
- Live Demo (4-5 minutes)
  - Start by wrapping up the full pipeline, from raw input how to get to the final answer.
  - The demo should be live not recorded video
  - We should be able to switch between the embedding model, and switch between the LLMs
  - Using your chosen questions to answers, we evaluate the integration of your pipeline
  - The demo must show *integration*. **not isolated components.**
  - The UI should be reflecting the process done in the background, check the Build UI section in the project description.

Faculty of Media Engineering and Technology
German University in Cairo
Dr. Nourhan Ehab

What not to do:

- Do NOT explain concepts that were explained in the lab
- Do NOT add introductions, problem statements, motivations, or related work/framework descriptions.
- In your slides, Do NOT depend only on diagrams, only on text, only on screenshots
- No text-heavy descriptions.Highlight only **what you actually implemented**.
- **Do NOT include dataset descriptions or high-level overviews of the theme.**

Regulations:

- Each team member is responsible for one component of the milestone. You will be required to indicate who is responsible for which component when scheduling your evaluation. (Make sure to distribute the work **equally** between team members (e.g. doing the slides of the presentation is NOT an individual task))
- All team members must participate in the presentation.
- Presentation slides must be submitted along with the milestone deliverables by **December 15th at 23:59**.
- You must work and present as a unified team with a fully integrated, functional project. (Randomly assigned teams included.)
- If one or more members are unable to attend or present, this should be communicated from the start of the evaluation time so that the remaining members may present their parts to avoid losing the deliverable grades.

Late Policy:

We respect your time, especially during revision week, and we expect you to respect other teams' time as well. Delays during MS1 caused significant disruptions to the schedule; this will **not** be tolerated in the MS3 evaluation.

- **Being 19 minutes late or more** will result in **cancellation of your evaluation** and a grade of **0** for both the deliverables and the evaluation.
- **Being 15-18 minutes late** will result in a **30% deduction** from your final grade.
- **Being 10-14 minutes late** will result in a **15% deduction** from your final grade.

Faculty of Media Engineering and Technology
German University in Cairo
Dr. Nourhan Ehab

Please keep in mind that your evaluation slot will end at the scheduled time. Arriving late reduces the number of rounds available for each team member and decreases your opportunity to correct any mistakes during the individual evaluation.