Knowledge Graphs with Large Language Models

MSc in AI and Data Science, 2025-2026

Panos Alexopoulos

Assignment 1: QA System and Knowledge Graph Specification

Introduction

The CEO of your company has just returned from an AI conference and is very excited about AI. They've heard about knowledge graphs and large language models, and they want to see something tangible. For that, they ask your team to develop a Proof-of-Concept (POC) system that demonstrates the value of these technologies for your company. However, they don't have a specific application in mind.

Given these parameters, your team decides to develop a small Question Answering (QA) system using the GraphRAG paradigm, a technique where an LLM uses a knowledge graph to answer questions.

You volunteer to lead the development, and your first step is to specify the requirements for the QA system and for the knowledge graph that will power it.

Tasks

Task 1: Select a scenario and a domain

Choose a realistic area where a GraphRAG-based QA system could demonstrate value (e.g., HR, product catalog, customer service, research management, healthcare). Take into consideration the availability of data that you will need to construct the knowledge graph (e.g. existing graphs, ontologies, databases, documents, etc).

Task 2: Describe the envisioned QA system

Write a short narrative describing the QA system you plan to build. Your description should clearly explain:

- **What the system does:** What kind of information or answers it provides.
- **Who the users are:** The intended audience (e.g., customer support agents, HR analysts, researchers, executives).
- **Why it is useful:** The problems it solves or the value it creates for the organization.

Task 3: Define the system's question space

Specify the kind and range of questions your QA system should be able to answer.

- Include 15–20 concrete example questions, phrased as users might ask them.

- Cover different levels of complexity:

    - Simple factual queries (e.g., "Who manages project X?"
    - Relational questions (e.g., "Which employees have worked on the same projects as Maria?")
    - Analytical or summarizing questions (e.g., "Which departments have the highest turnover?")
    - Any other type of question you might see fit.

Task 4: Derive requirements§ for the knowledge graph

Define questions and/or other requirements that express what the knowledge graph must contain in order to support the answering of task 3 questions.

Deliverables

A short report describing the outcomes of the 4 tasks.

Important notice

You are allowed to use an LLM to help you with the above tasks. However, if you do that, you need to be critical against the suggestions that the LLM might give you and correct/adapt them as needed. **Remember that you are solely responsible for the end result.**