LLM Reasoning & Personality Knowledge Graph Challenge

Summary

While most people now use LLMs for code generation, not everybody actually understands the topics. In this exercise, you are *expected* to leverage an LLM as a research and guidance tool during this assessment. The primary objective is to evaluate your ability to conduct independent research and effectively use LLMs for code assistance — *not* merely to write code. (Any hands-on coding can be discussed during a follow-up interview.)

Deliverables

- Public GitHub repository containing the code (either Pure Python code or notebook is fine)
- Short report explaining design choices, insights, and limitations. You may use LLM to help for the generation.
- Your shared LLM session (we want to see your thinking process)

Problem Statement

Develop a Python solution that takes a text document, extract and build a knowledge graph (KG) from the document. The graph will represent the knowledge from the document in a structured way. Furthermore, the personality of the subjects are also modelled in the graph.

This is an open problem, there is no fixed requirement on the input sources, modeling approach, and outputs. The most important thing is your logical thinking in solving the problem. You will need to consult an LLM and share with us the session.

1. **Research with an LLM**
    - Use your preferred LLM to conduct quick background research on the topic.
    - Make sure you ***understand*** *the* problem. Ask LLM for anything that you don't understand. Some of the possible questions you might want to ask):
        - What is a knowledge graph, and what is its purpose?
        - How can personality traits be represented or connected within such a graph?
    - The LLM will provide multiple approaches — you need to understand the topic to **justify your chosen design** in your report.
2. **Code Generation & Experimentation**
    - Ask the LLM to generate code snippets that help you implement your ideas.
    - Run and test these snippets yourself.

- Iterate with the chatbot to refine your design and improve the model (everything from inputs, data representation, outputs etc).
- Of course, you can do some manual adjustments. Feel free to do it.

3. **Documentation**
    - Share your LLM session (e.g., use the “Share” feature in ChatGPT or provide screenshots) with us.
    - Clearly explain the reasoning behind your chosen approach in a report. You should ask the LLM to generate the report, don’t manually write it yourself.

In your solution, please consider the following. Everything can come from LLM:

- How and what synthetic data to generate (you won’t have real data, so you will need to generate synthetic data)? Can you justify your decisions?
- How to evaluate the model and what metrics to use? Can you justify your decisions?
- How to do the implementation? How to represent the personality? Can you justify your decisions?
- How to pipe the LLM workflows (remember KG graph construction is *not* a single prompt, it should be a chain of prompts). Ask your LLM. Can you justify your decisions?
- How to do your data processing (e.g. what normalisation approach do you want to use, what do you want to normalise, what do you don’t want to normalise?). Can you justify your decisions?