Pb.03 – Complex prompting strategies

1. Introduction to this lab

In this lab, you will work in a Colab environment. Using Ollama we install two large language models (LLMs): DeepSeek R1 and Llama 3.2. Both of those models were released in 2025 as state-of-the-art, open-weight models. Open-weight means the weights of the model (learned parameters) have been made available to the public to download after the costly training.

You can get started using this link:
https://colab.research.google.com/github/benikm91/zhaw_ai2_lab3/blob/main/planning_with_reasoning_llm.ipynb

2. Demonstrating complex prompting using state-of-the-art LLMs

There are four consecutive tasks to solve in the Colab notebook, all showing you how to prompt state-of-the-art LLMs with increasing complexity. The notebook contains step-by-step descriptions for each task. Here's a brief overview:

- **Task 1** sets up the notebook, installs both models and tests them using basic prompting.
- **Task 2** introduces zero-shot tool usage with LLMs using Llama 3.2.
- **Task 3** demonstrates an agentic workflow. The R1 reasoning model is the planner: Given a description of available tools, it creates an abstract plan to solve a user prompt. Then, the Llama 3.2 model is the plan executer. Prompted with R1's plan, the available tools, and their API, it generates tool calls to solve the user prompt.
- **Task 4** is the open-ended challenge. Now it's your turn! Apply the concepts and techniques you've learned in Tasks 1 through 3 to tackle a new scenario - think of it as a problem you might encounter or define yourself