

Task 1: Generation of design documentation

The essence of the problem

Develop a system that automatically generates structured design documentation based on a set of disparate materials using a language model.

In real-world engineering and research projects, source data is rarely presented as a complete document. Information is dispersed across various sources: notes, code snippets, technical descriptions, budgets, and discussions. The system's task is to collect this data, resolve inconsistencies, and produce a coherent document.

The system must:

- extract facts from unstructured sources;
- combine information from different files;
- preserve the meaning without distortion;
- generate a document suitable for use (application, specifications, project description);
- minimize hallucinations.

In fact, the problem being solved is

the construction of a reliable document generation pipeline from disparate data.

Input data

Formats:

- .txt , .md , .json , .py , .java

Data can be:

- partially duplicated;
- contradictory;
- distributed across different files;
- unstructured.

Requirements for the solution

- use of LLM;
- HTTP API:

POST /generate_document

- strict structure of the result;
- reproducibility;
- independence from a specific data set.

Result format

```
{
  "document": {
    "project_overview": "",
    "goals": [],
    "requirements": [],
    "technical_solution": "",
    "architecture": "",
    "team": [],
    "timeline": "",
    "budget": "",
    "risks": []
  },
  "metadata": {
    "model_name": "",
    "llm_calls": 0,
    "total_tokens": 0,
    "duration_ms": 0
  },
  "trace": {
    "steps": []
  }
}
```

Evaluation criteria

- Accuracy and completeness - 40%
- Absence of hallucinations - 20%
- Efficiency (tokens) - 10%
- LLM call rate is 10%
- Completion time - 10%
- Structure and format - 10%

Expected result

A system capable of reliably generating accurate design documentation from previously unknown datasets. Improvements that enhance the usability and practical value of the solution are welcome.