

Задание 1: Генерация проектной документации

Суть задачи

Разработать систему, которая на основе набора разрозненных материалов автоматически формирует структурированную проектную документацию с использованием языковой модели.

В реальных инженерных и исследовательских проектах исходные данные редко представлены в виде готового документа. Информация распределена по различным источникам: заметки, фрагменты кода, технические описания, бюджеты, обсуждения. Задача системы — собрать эти данные, устранить противоречия и сформировать цельный документ.

Система должна:

- извлекать факты из неструктурированных источников;
- объединять информацию из разных файлов;
- сохранять смысл без искажения;
- формировать документ, пригодный для использования (заявка, ТЗ, описание проекта);
- минимизировать галлюцинации.

Фактически решается задача:

построения надёжного pipeline генерации документа из разрозненных данных.

Входные данные

Форматы:

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

Данные могут быть:

- частично дублирующимися;
- противоречивыми;
- распределёнными по разным файлам;
- неструктурированными.

Требования к решению

- использование LLM;
- HTTP API:

POST /generate_document

- строгая структура результата;
- воспроизводимость;
- независимость от конкретного набора данных.

Формат результата

```
{
  "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": []
  }
}
```

Критерии оценки

- Достоверность и полнота — 40%
- Отсутствие галлюцинаций — 20%
- Эффективность (токены) — 10%
- Количество вызовов LLM — 10%
- Время выполнения — 10%
- Структура и формат — 10%

Ожидаемый результат

Система, способная устойчиво формировать корректную проектную документацию из ранее неизвестных наборов данных. Приветствуются доработки, повышающие удобство и практическую ценность решения.