AI for creating Standardized Reports

Graduate

Kailing Peng

Introduction: This thesis focuses on the intelligent generation of standardized reports - specifically the 8D report, a widely adopted framework in quality management used to systematically identify root causes, implement corrective actions, and ensure continuous improvement in response to customer complaints or internal quality issues.

Approach: Rather than fine-tuning a dedicated model for each specific report template, this thesis investigates the use of general-purpose AI models - specifically the large language model (LLM) GPT-4o-mini - to interpret report structures and task specifications through prompt engineering and Retrieval-Augmented Generation (RAG). The system aims to retrieve relevant information from heterogeneous data sources and accurately fill the required report fields, while maintaining constraints such as report structure, data type consistency, object integrity, and completeness. Key challenges include:

- Secure handling of sensitive and business-critical data
- Understanding and processing unstructured information from diverse sources
- Supporting flexible, customizable report formats
- Minimizing LLM hallucinations and ensuring factual correctness
- Integrating human-in-the-loop validation for final quality assurance

To address these challenges, an LLM-powered pipeline is developed that includes input document preprocessing, data vectorization, and structured report content generation. The pipeline also features an automated self-reflection step that guides the user in addressing missing or suspicious field values before intelligently updating the report based on user instructions. Various strategies - such as few-shot prompting, static and dynamic query generation, multi-turn interactions, and ReAct-style agents - are evaluated to identify optimal configurations. Several open-source frameworks are assessed, with particular focus on the toolkit Docling, valued for its ability to handle multiple document formats, perform advanced PDF parsing and Optical Character Recognition (OCR), and support integrated RAG workflows for automated report generation.

Result: The proposed generative AI-based pipeline is evaluated for reliability, efficiency, cost, and scalability. Results show that high-quality standardized report template creation and content generation can be achieved without relying on rigid, code-based applications. Based on the observations and research, the thesis proposes different combinations of tools and engineering strategies to achieve an optimal cost-performance balance, depending on task complexity and input structure.

The final prototype demonstrates that, with thoughtful system design, AI-driven report generation is both practical and adaptable for real-world business use.
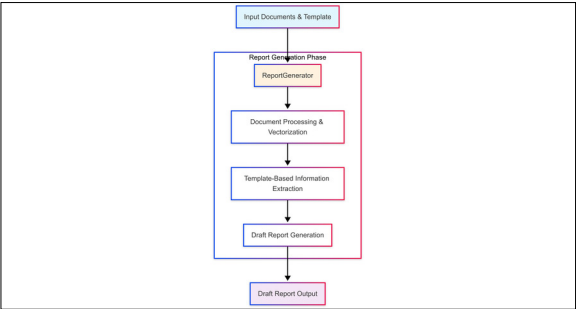
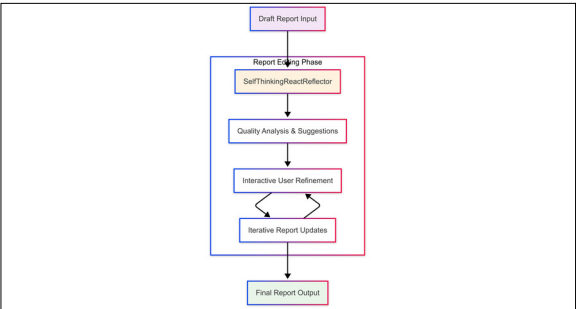**Report Template Creation from Learning**
Own presentment
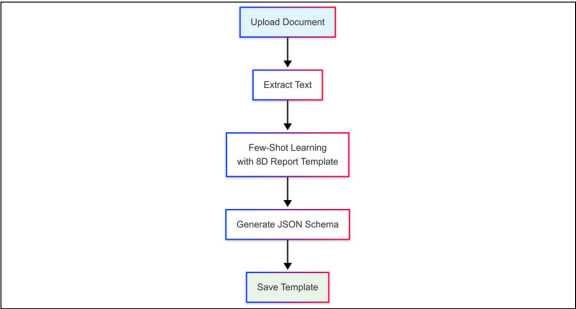
Report Generation Automatized Processing
Own presentment

Report Post-editing in Chat
Own presentment

Advisor
Prof. Dr. Mitra Purandare

Co-Examiner
Fabian Kostadinov

Subject Area
Artificial Intelligence

Project Partner
Mettler Toledo