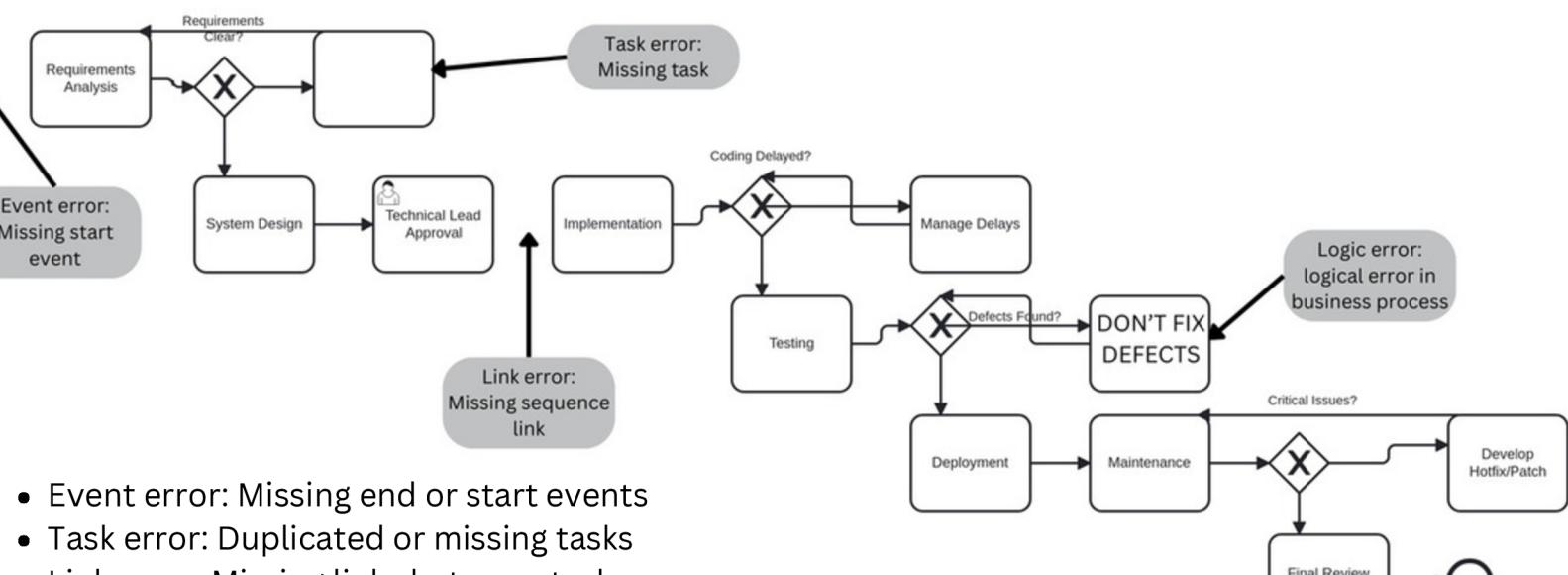


Introduction

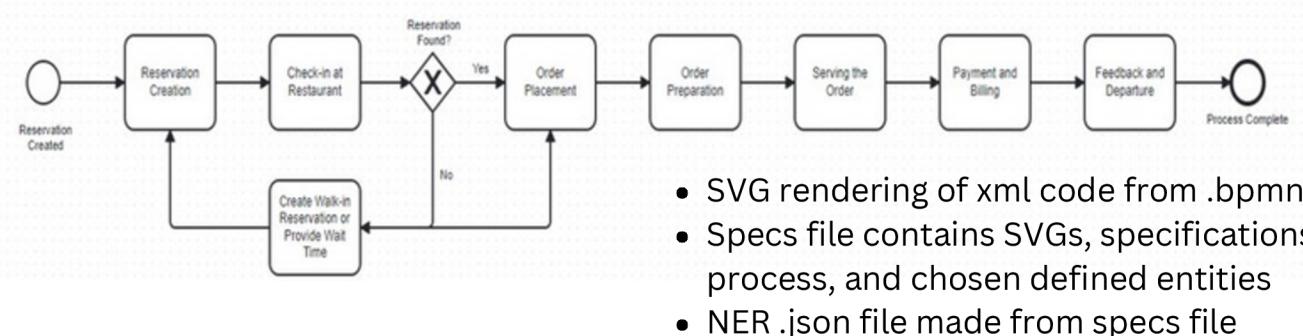
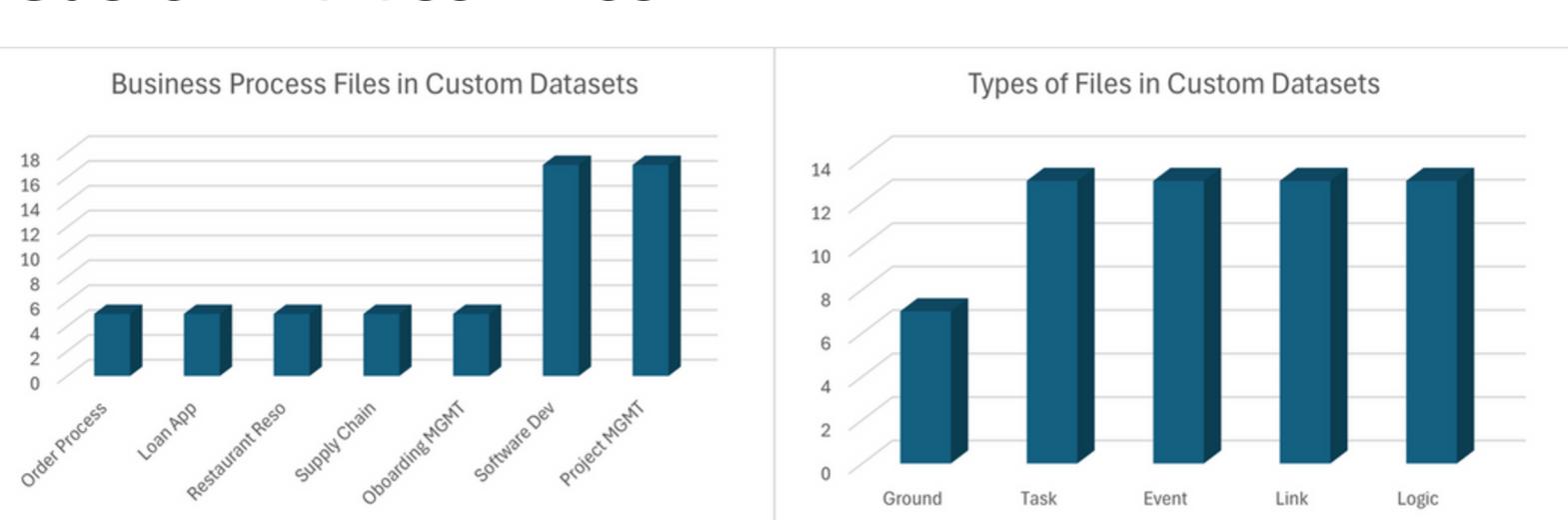
Business Process Logic (BPL) systems are critical yet prone to faults that disrupt workflows and escalate costs. This project leverages cutting-edge Large Language Models (LLMs) like GPT-4o, combined with advanced NLP tools such as pdfplumber, spaCy, and SentenceTransformer, to detect and resolve logical discrepancies in BPL systems with expert precision. By automating fault detection and integrating continuous learning, enabling businesses to achieve unparalleled efficiency, accuracy, and operational excellence.

Types of errors

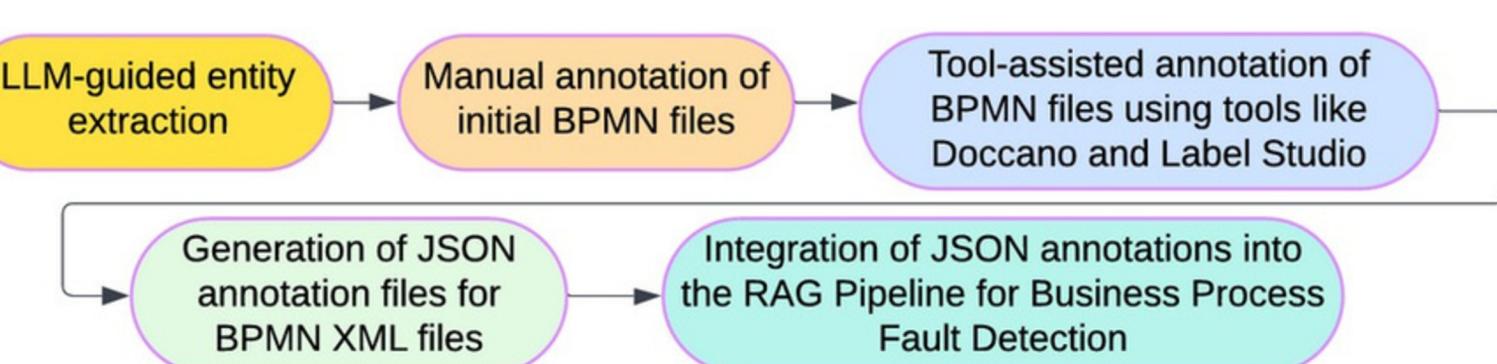


- Event error: Missing end or start events
- Task error: Duplicated or missing tasks
- Link error: Missing links between tasks
- Logic error: Process not following instructions

Custom Dataset Files

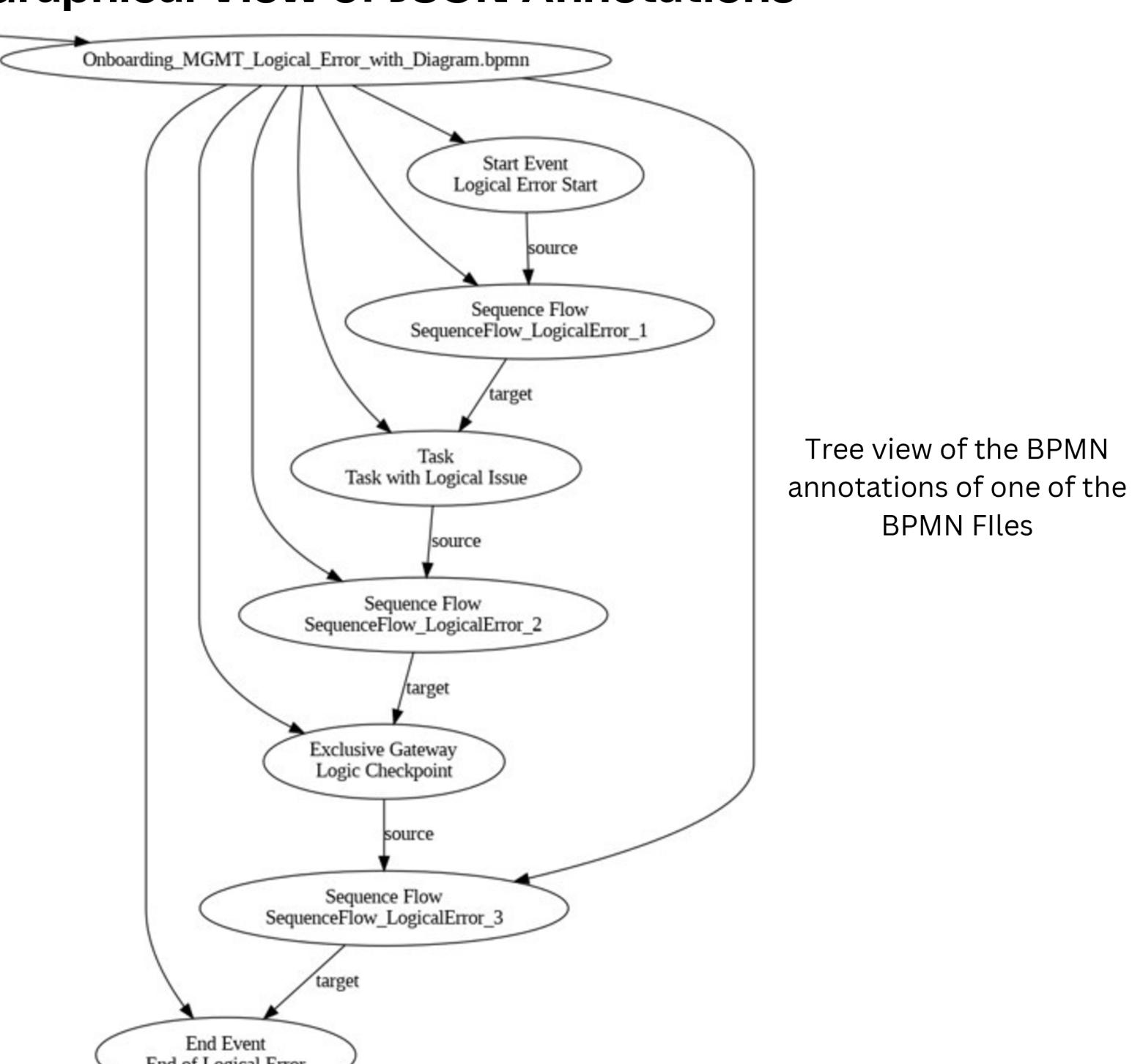


NER Flowchart



Annotation workflow: LLM-assisted extraction, manual refinement, tool-supported annotation, JSON generation, and integration into a RAG pipeline for fault detection.

Graphical View of JSON Annotations

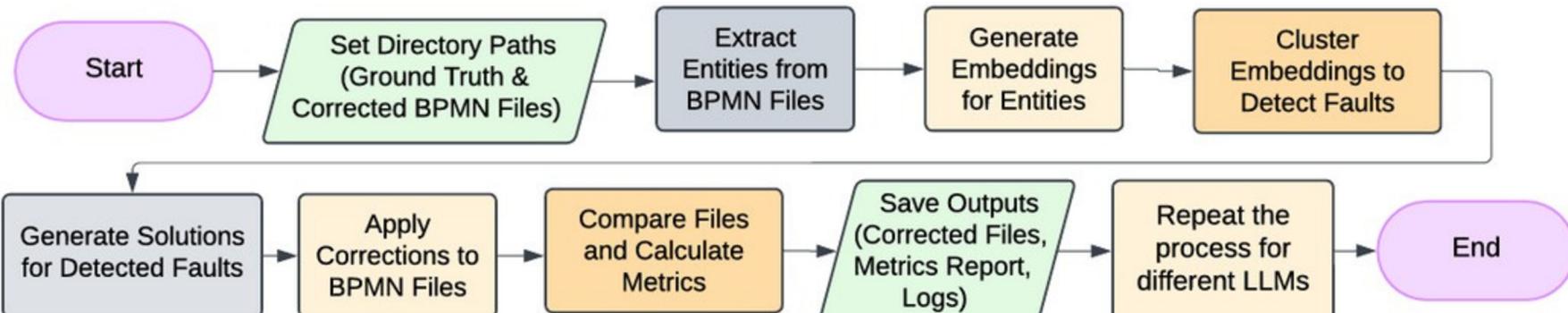


Tree view of the BPMN annotations of one of the BPMN files

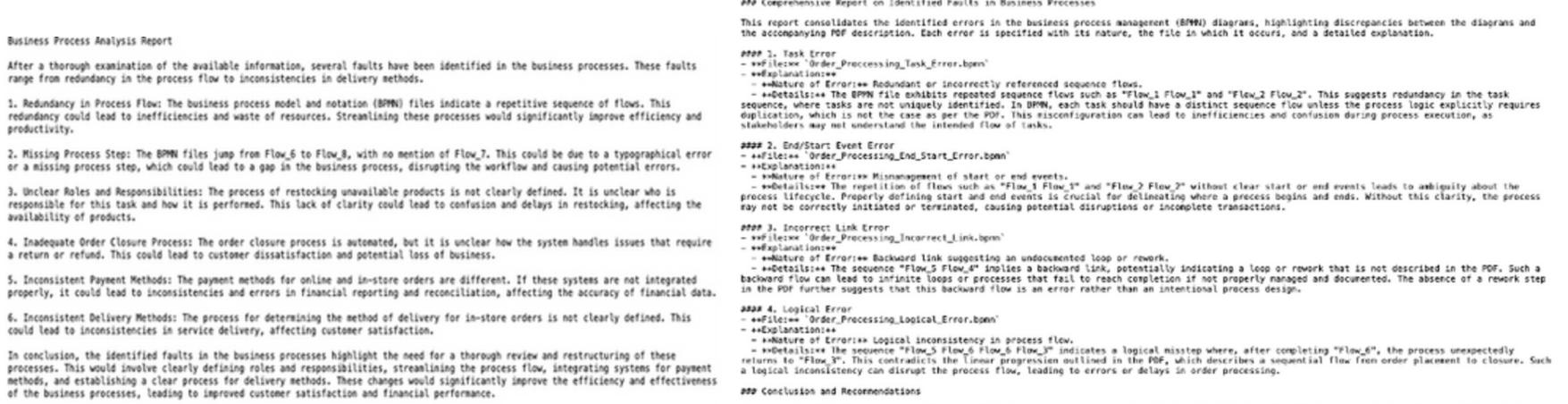
The Role of RAG Pipeline in BPMN Fault Detection

- Identifies structural and logical anomalies by combining retrieval of BPMN rules with generative reasoning.
- Uses external knowledge for precise, science-driven fault resolutions.
- Efficiently processes large BPMN repositories with targeted rule-based retrieval and corrections.
- Resolves ambiguities by grounding generative outputs in explicit domain knowledge.

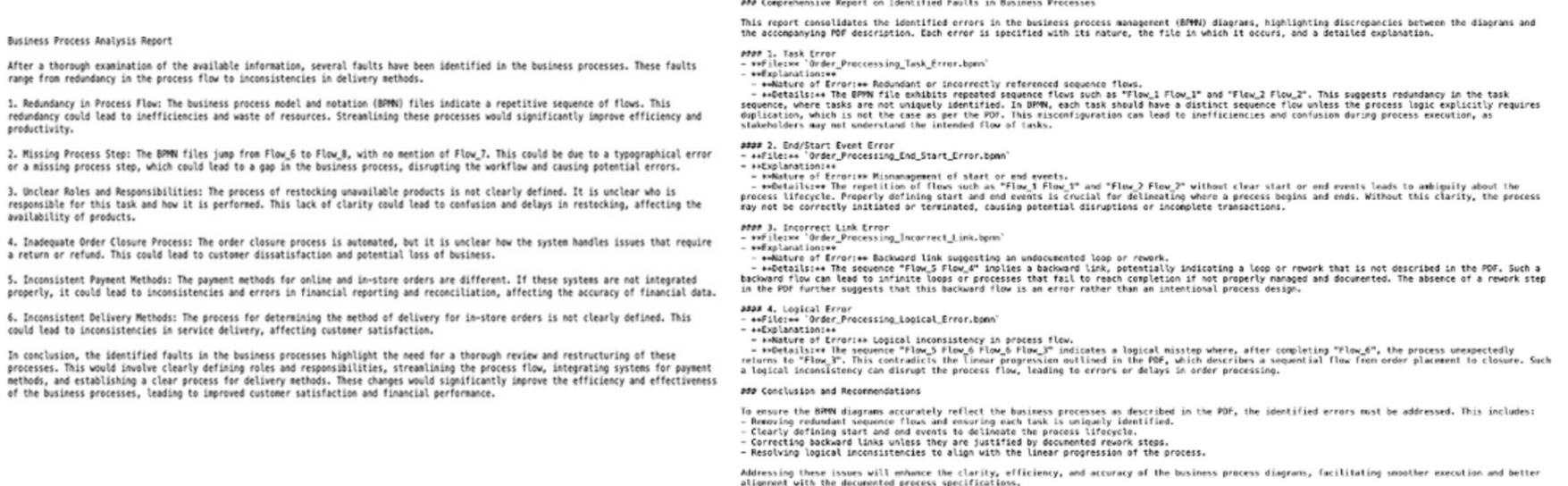
RAG Pipeline flowchart



Before RAG



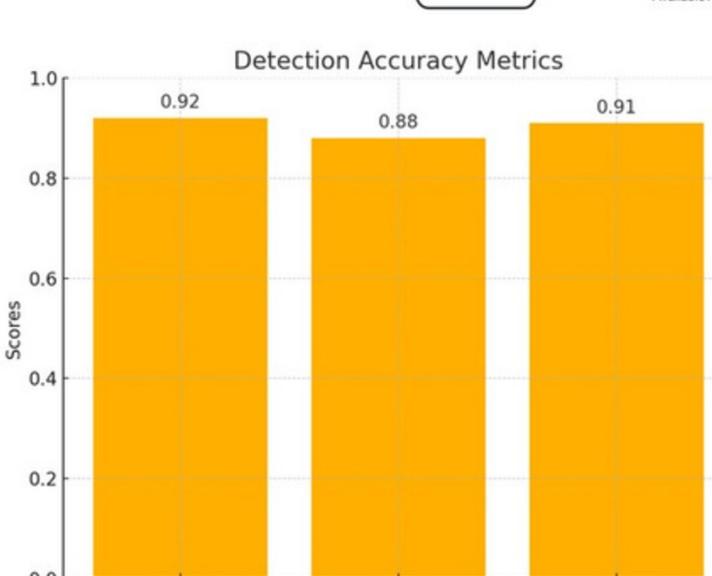
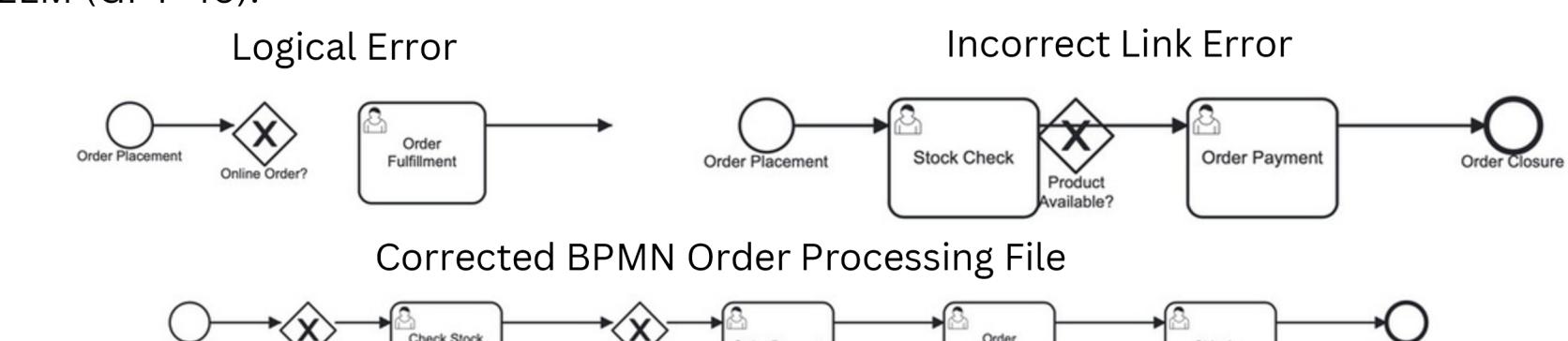
After RAG



- Before RAG Output:** Broad fault descriptions with limited specificity and no actionable insights.
- After RAG Output:** Precise fault identification with detailed explanations and exact BPMN file locations.

Results

Below are the images of the errors identified in the Order Processing business process file, followed by the corrected Order Processing business process file generated by the LLM (GPT-4o).



These results showcase clear visualizations of errors like logical mismatches and incorrect links, alongside a corrected BPMN workflow that eliminates these issues. High detection accuracy metrics—Precision (0.92), Recall (0.88), and F1 Score (0.91)—underscore the system's effectiveness, making it a game-changer for streamlining business processes.

Future Directions

- Enhancing Accuracy:** Boost fault detection precision, especially for complex and nuanced errors.
- Refining Solutions:** Deliver context-aware, automated resolutions to minimize manual intervention.
- Expanding NER:** Broaden entity coverage for comprehensive workflow analysis.
- Testing on LLMs:** Benchmark performance across diverse models to identify the best fit.
- Adaptive Learning:** Integrate feedback-driven learning for continuous system improvement.
- Industry Scaling:** Tailor solutions to diverse workflows for seamless deployment across industries.