Fault Detection in Business Process Logic Systems Using LLMs

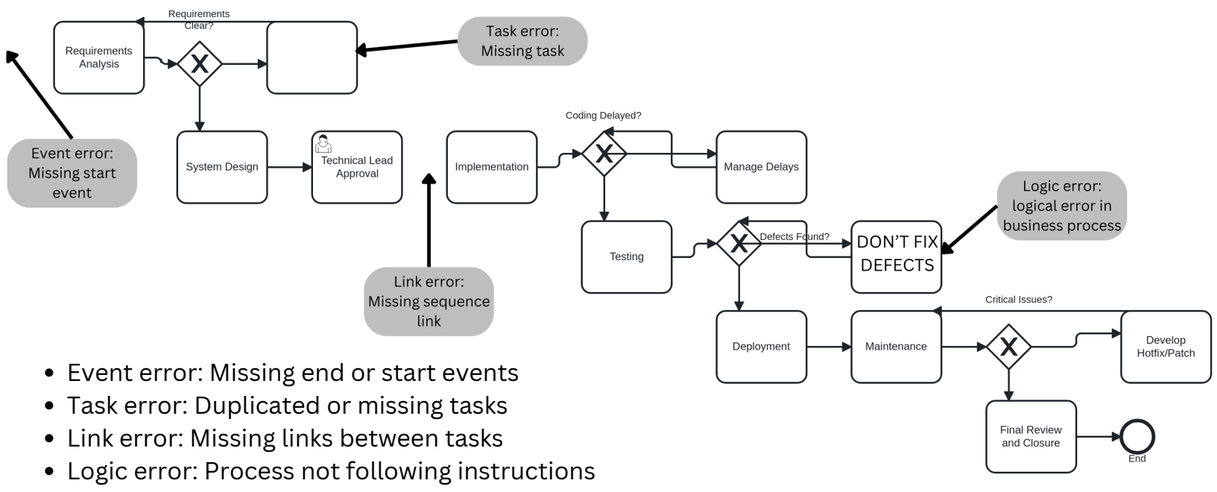
CS 3892/5892: Project in Data-Centric AI and Mining

Sunny Bhatt Srikanth Narayanan Josue Garcia Flores Mamatha Reddy Tamilisetti

Introduction

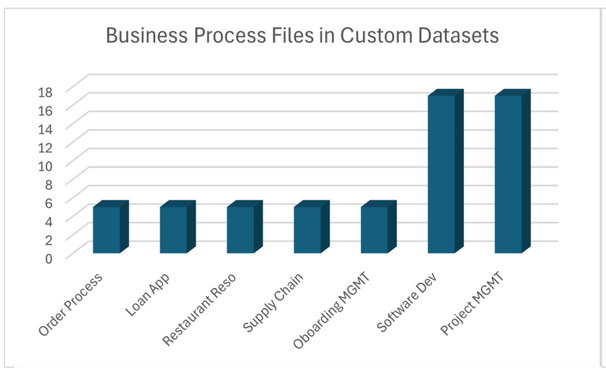
Business Process Logic (BPL) systems are critical yet prone to faults that disrupt workflows and escalate costs. This project leverages cutting-edge Large Language Models (LLMs) like GPT-4o, combined with advanced NLP tools such as pdfplumber, spaCy, and SentenceTransformer, to detect and resolve logical 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.

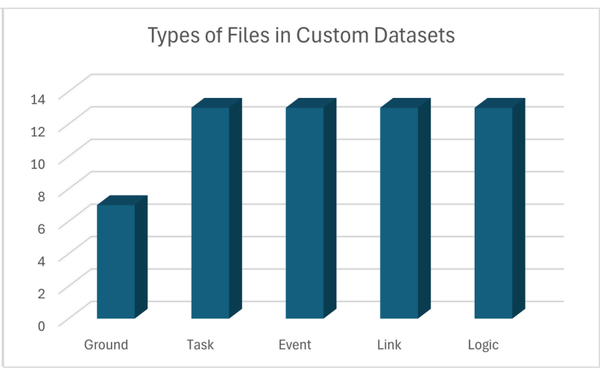
Types of errors

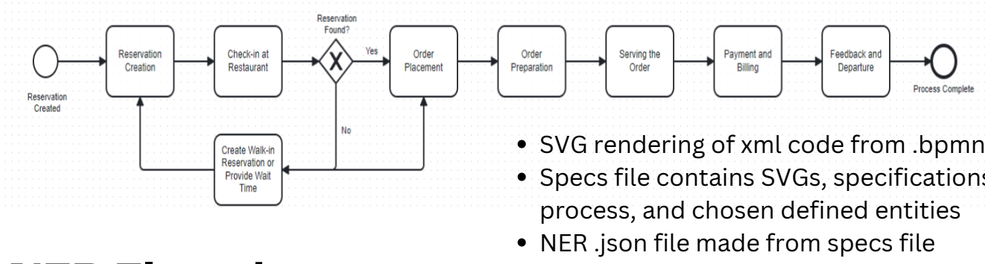


- Event error: Missing end or start events
- Task error: Duplicated or missing tasks
- Link error: Missing links between tasks
- Logic error: Process not following instructions

Custom Dataset Files

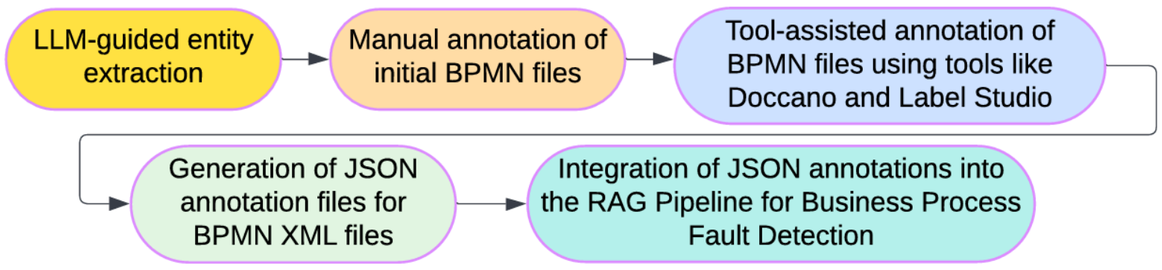






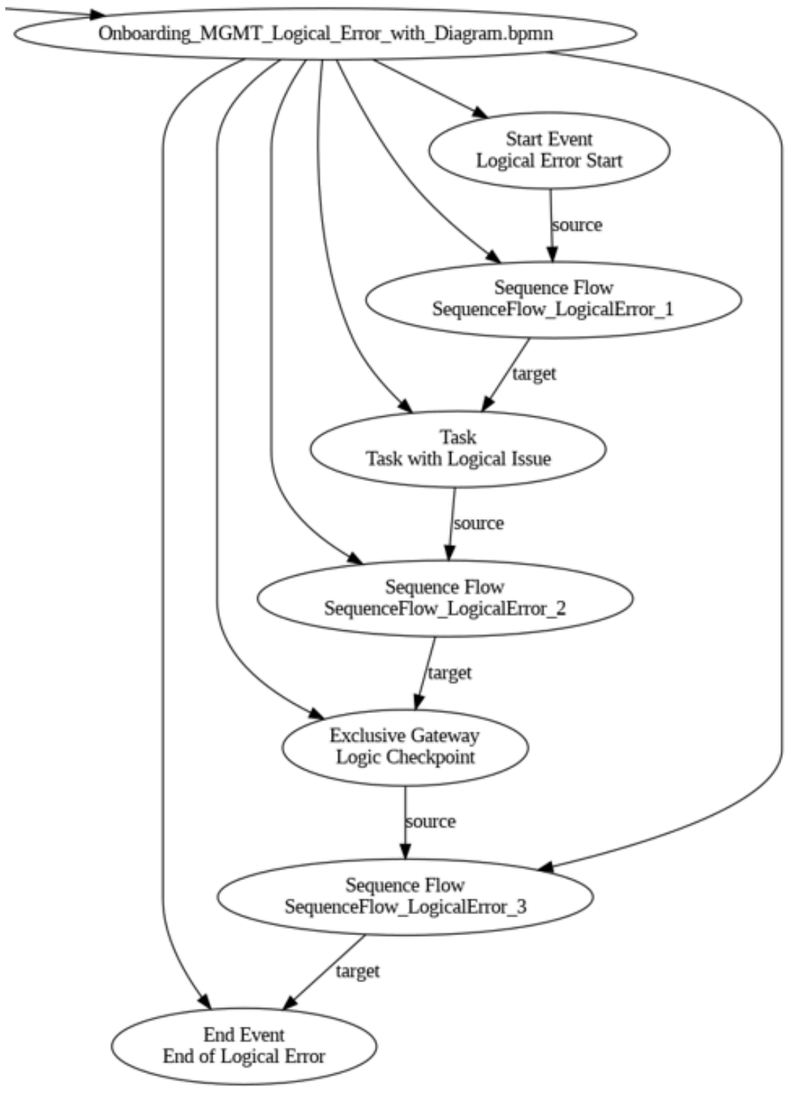
- SVG rendering of xml code from .bpmn file
- Specs file contains SVGs, specifications of business process, and chosen defined entities
- NER .json file made from specs file

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

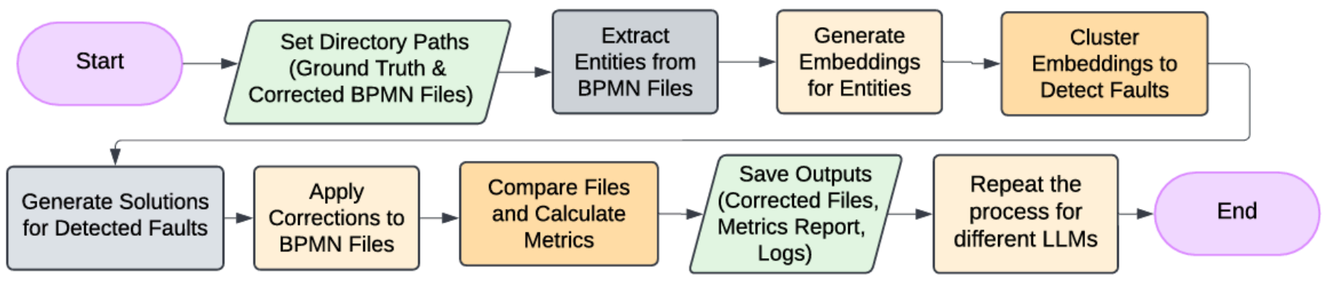


Tree view of the BPMN annotations of one of the BPMN FIles

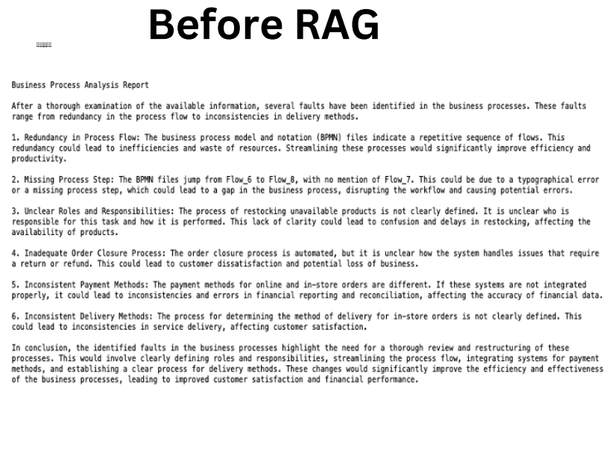
The Role of RAG Pipeline in BPMN Fault Detection

- 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



Business Process Analysis Report

After a thorough examination of the available information, several faults have been identified in the business processes. These faults range from redundancy in the process flow to inconsistencies in delivery methods.

1. Redundancy in Process Flow: The business process model and notation (BPMN) files indicate a repetitive sequence of flows. This redundancy could lead to inefficiencies and waste of resources. Streamlining these processes would significantly improve efficiency and productivity.

2. Missing Process Step: The BPMN files jump from Flow_6 to Flow_8, with no mention of Flow_7. This could be due to a typographical error or a missing process step, which could lead to a gap in the business process, disrupting the workflow and causing potential errors.

3. Unclear Roles and Responsibilities: The process of restocking unavailable products is not clearly defined. It is unclear who is responsible for this task and how it is performed. This lack of clarity could lead to confusion and delays in restocking, affecting the availability of products.

4. Inadequate Order Closure Process: The order closure process is automated, but it is unclear how the system handles issues that require a return or refund. This could lead to customer dissatisfaction and potential loss of business.

5. Inconsistent Payment Methods: The payment methods for online and in-store orders are different. If these systems are not integrated properly, it could lead to inconsistencies and errors in financial reporting and reconciliation, affecting the accuracy of financial data.

6. Inconsistent Delivery Methods: The process for determining the method of delivery for in-store orders is not clearly defined. This could lead to inconsistencies in service delivery, affecting customer satisfaction.

In conclusion, the identified faults in the business processes highlight the need for a thorough review and restructuring of these processes. This would involve clearly defining roles and responsibilities, streamlining the process flow, integrating systems for payment methods, and establishing a clear process for delivery methods. These changes would significantly improve the efficiency and effectiveness of the business processes, leading to improved customer satisfaction and financial performance.

After RAG

Comprehensive Report on Identified Faults in Business Processes

This report consolidates the identified errors in the business process management (BPMN) diagrams, highlighting discrepancies between the diagrams and the accompanying PDF description. Each error is specified with its nature, the file in which it occurs, and a detailed explanation.

1. Task Error
- **File:** `Order_Proccessing_Task_Error.bpmn`
- **Explanation:**
  - **Nature of Error:** Redundant or incorrectly referenced sequence flows.
  - **Details:** The BPMN file exhibits repeated sequence flows such as "Flow_1 Flow_1" and "Flow_2 Flow_2". This suggests redundancy in the task sequence, where tasks are not uniquely identified. In BPMN, each task should have a distinct sequence flow unless the process logic explicitly requires duplication, which is not the case as per the PDF. This misconfiguration can lead to inefficiencies and confusion during process execution, as stakeholders may not understand the intended flow of tasks.

2. End/Start Event Error
- **File:** `Order_Processing_End_Start_Error.bpmn`
- **Explanation:**
  - **Nature of Error:** Mismanagement of start or end events.
  - **Details:** The repetition of flows such as "Flow_1 Flow_1" and "Flow_2 Flow_2" without clear start or end events leads to ambiguity about the process lifecycle. Properly defining start and end events is crucial for delineating where a process begins and ends. Without this clarity, the process may not be correctly initiated or terminated, causing potential disruptions or incomplete transactions.

3. Incorrect Link Error
- **File:** `Order_Processing_Incorrect_Link.bpmn`
- **Explanation:**
  - **Nature of Error:** Backward link suggesting an undocumented loop or rework.
  - **Details:** The sequence "Flow_5 Flow_4" implies a backward link, potentially indicating a loop or rework that is not described in the PDF. Such a backward flow can lead to infinite loops or processes that fail to reach completion if not properly managed and documented. The absence of a rework step in the PDF further suggests that this backward flow is an error rather than an intentional process design.

4. Logical Error
- **File:** `Order_Processing_Logical_Error.bpmn`
- **Explanation:**
  - **Nature of Error:** Logical inconsistency in process flow.
  - **Details:** The sequence "Flow_5 Flow_6 Flow_6 Flow_3" indicates a logical misstep where, after completing "Flow_6", the process unexpectedly returns to "Flow_3". This contradicts the linear progression outlined in the PDF, which describes a sequential flow from order placement to closure. Such a logical inconsistency can disrupt the process flow, leading to errors or delays in order processing.

Conclusion and Recommendations

To ensure the BPMN diagrams accurately reflect the business processes as described in the PDF, the identified errors must be addressed. This includes:
- Removing redundant sequence flows and ensuring each task is uniquely identified.
- Clearly defining start and end events to delineate the process lifecycle.
- Correcting backward links unless they are justified by documented rework steps.
- Resolving logical inconsistencies to align with the linear progression of the process.

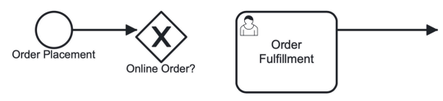
Addressing these issues will enhance the clarity, efficiency, and accuracy of the business process diagrams, facilitating smoother execution and better alignment with the documented process specifications.

- 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

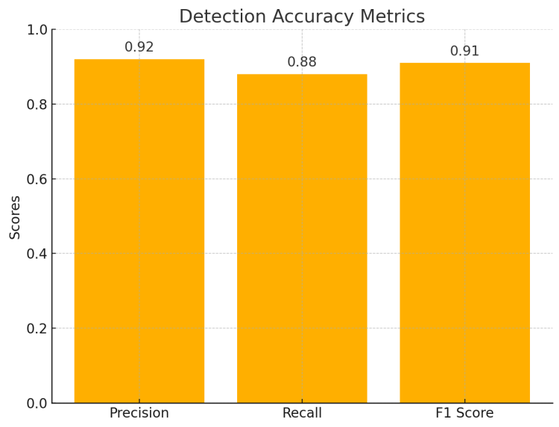
Below are the images of the errors identified in the Order Processing business process file, followed by the corrected Order Processing business process file generated by the LLM (GPT-4o).

Logical Error



Incorrect Link Error

Corrected BPMN Order Processing File



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 Solution**s: 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.