# Leveraging LLMs for Prescriptive Business Process Monitoring and its Evaluation for a Real-world Data Set

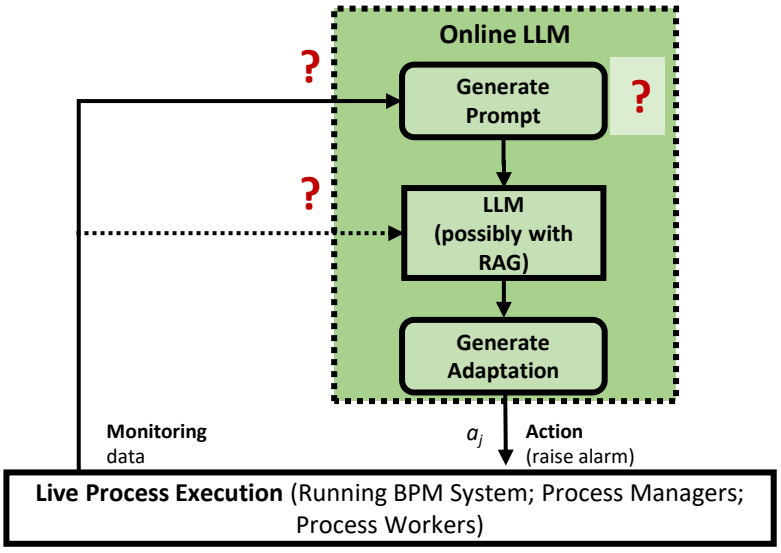
**BA Topic Description**



Prescriptive business process monitoring provides decision support to process managers on when and how to adapt an ongoing business process to prevent or mitigate an undesired process outcome. A key challenge is balancing prediction accuracy with earliness: earlier predictions allow more adaptation time, but are often less reliable. Different approaches were presented in the literature to reconcile this trade-off. These approaches include using a static prediction point, applying empirical thresholding, and using online reinforcement learning. In our recent journal article, we evaluated and compared the performance of these different approaches [Metzger et al., 2023 @ Information Systems; https://doi.org/10.1016/j.is.2023.102254].

The aim of the bachelor thesis is to build on these previous results and existing BPM benchmark data sets. In particular the thesis shall leverage Large Language Models (LLMs), as a powerful generative AI approach, for prescriptive business process monitoring. LLMs have shown an impressive improvement in performance and are being applied to an increasingly broader range of problems. As an example, in the field of adaptive software systems, they were proposed for generating (new) adaptations at run-time [Li et al., 2024 @ TAAS; https://doi.org/10.1145/3686803].

The figure below depicts the high-level architecture of the envisioned solution of the BA thesis.




The main idea is to transfer the design-time usage of LLMs to run-time. Key research questions include “how to engineer the inputs for the LLM?” (in particular, how to convert the process monitoring data into specific prompts and/or RAG inputs), as well as “how to parametrize the LLM?” (e.g., to cope with hallucinations).

To assess the effectiveness of the LLM solution it shall be evaluated using one of the following four real-world BPM data sets as used in following [Metzger et al., 2023]:

Data sets used in experiments.

| Name | Type of deviation | Rate of deviations | Size of "Test" data | Process variants | Max. prefix length |
|---|---|---|---|---|---|
| BPIC12 | Unsuccessful loan application | 25% | 4,361 | 3587 | 48 |
| BPIC17 | Unsuccessful loan application | 41% | 10,500 | 2087 | 71 |
| Traffic | Unpaid traffic fines | 58% | 50,117 | 185 | 5 |
| Cargo | Delay in cargo delivery | 31% | 1,313 | 144 | 21 |