

Note: All topics are available to both bachelor's and master's students.

Topic 1: Object-Centric Process Mining

The goal of process mining is to gain insights from event logs that record real-life process executions. In traditional process mining, these event logs have a single case notion that determines what a process instance is. For example, if the case notion is a sales order, then all events related to “Order 1” are assumed to belong to the same instance of a process. However, in real-life processes, events typically relate to several entities at the same time – for instance, an order may include multiple packages and items and be associated with a customer. “Flattening” this interplay of entities by assigning events to a single case notion provides only a limited view on the real process and can lead to convergence and divergence issues in the resulting event log. Object-centric process mining aims to address this by getting rid of the single case notion and developing new event log formats and process mining techniques.

The objective of this seminar thesis is to compile a high-level overview of the current state of the art in object-centric process mining. This involves identifying, for example, object-centric data formats and algorithms for process discovery or conformance checking and analyzing how they differ from traditional ones. Based on the results of this review, the participant should then pinpoint research gaps, i.e., determine for which aspects of process mining there are currently no object-centric counterparts or alternatives.

Literature:

- [1] Van der Aalst, W.M.P. (2019). [Object-Centric Process Mining: Dealing with Divergence and Convergence in Event Data](#). Software Engineering and Formal Methods, 3-25.
- [2] Ghahfarokhi, A.F., Park, G., Berti, A., & van der Aalst, W.M.P. (2021). [OCEL: A Standard for Object-Centric Event Logs](#). New Trends in Database and Information Systems, 169-175.

Topic 2: Extracting Process Information from Natural Language

Business Process Management (BPM) aims to improve organizational activities and their outcomes by managing the underlying processes. To achieve this, it is often necessary to consider information from various sources, including unstructured textual documents. Therefore, researchers have developed several BPM-specific solutions that extract information from textual documents using Natural Language Processing techniques. Examples include deriving imperative or declarative process models from textual process descriptions or identifying candidate tasks for Robotic Process Automation.

The goal of this seminar thesis is to identify existing BPM approaches that make use of textual data. A particular focus should be put on finding publications (including pre-prints) that leverage Large Language Models and determining how their performance compares to more traditional approaches.

Literature:

- [1] Gross, M., Abb, L., Elsayed, N., & Rehse, J.R. (2023): [Large Language Models can accomplish Business Process Management Tasks](#). arXiv preprint arXiv:2307.09923.
- [2] Friedrich, F., Mendling, J., & Puhlmann, F. (2011): [Process Model Generation from Natural Language Text](#). International Conference on Advanced Information Systems Engineering.
- [3] Van der Aa, H., Di Ciccio, C., Leopold, H., & Reijers, H.A. (2019). [Extracting Declarative Process Models from Natural Language](#). International Conference on Advanced Information Systems Engineering.

Topic 3: Sustainability in Business Process Management

Organizations are increasingly focused on becoming more sustainable. This also extends to their operations, which means that there is an increasing need for sustainable business process management (BPM). Sustainable BPM aims to integrate environmentally and socially responsible practices into business processes. By relying on it, organizations can optimize operations, reduce waste, and promote long-term viability while minimizing the negative impact on the environment and society. Sustainable BPM involves identifying eco-friendly alternatives, promoting energy efficiency, and fostering responsible decision-making to create a positive impact on both the business and the broader community.

The goal of this seminar thesis is to provide an overview of concepts and methods of sustainable BPM. This should cover the development from “Green BPM” efforts from the 2010s all the way to the application of process mining for sustainability, which is researched today. Besides providing this historical perspective, the participant can choose whether to take a more management-focused approach, diving deeper into the required management practices to make BPM more sustainable, or a more technology-focused approach, diving deeper into the technical solutions that make sustainable BPM possible.

Literature:

- [1] Couckuyt, D., & Van Looy, A. (2020). [A systematic review of green business process management](#). Business Process Management Journal, 26(2), 421-446.
- [2] Stolze, C., Semmler, G., & Thomas, O. (2012). [Sustainability in business process management research—a literature review](#). American Conference on Information Systems.
- [3] Brehm, L., Slamka, J., & Nickmann, A. (2022). [Process Mining for Carbon Accounting: An Analysis of Requirements and Potentials](#). Digitalization Across Organizational Levels: New Frontiers for Information Systems Research, 209-244.

Topic 4: Visual Analytics Approaches for Sequential Data

Process data is typically represented as sequences of events that have a common identifier. In order to gain insights for process managers, the data should be visualized in an appropriate manner. To align process mining approaches with the visual needs of process managers, research can consider approaches from the domain of Visual Analytics to adjust their visualizations.

The goal of this seminar is to exploratively search for Visual Analytics approaches that are tailored to sequential data like process data. The participant should present different approaches, identify commonalities, and also elaborate on differences to the visualization in comparison to other data than sequential data. Further, the participant should analyze whether these approaches are applicable for process mining.

Since the number of publications in this field is expected to be large, a specification will be done by the participant in cooperation with the supervisor before the first milestone.

Literature:

- [1] Gotz, D., & Stavropoulos, H. (2014). [DecisionFlow: Visual Analytics for High-Dimensional Temporal Event Sequence Data](#). IEEE Transactions on Visualization and Computer Graphics 20(12), 1783-1792.
- [2] Kwon, B.C., Verma, J., & Perer, A. (2016). [Peekquence: Visual analytics for event sequence data](#). ACM SIGKDD Workshop on Interactive Data Exploration and Analytics.
- [3] Aigner, Wolfgang, et al. (2011) [Visualization of time-oriented data](#). Vol. 4.

Topic 5: The Influence of Sampling on Process Mining Results

Process mining involves analyzing event logs that capture the behavior of business processes within enterprise information systems. For large systems with frequently executed processes, these event logs can become computationally expensive to handle, especially for techniques like alignment-based conformance checking. To overcome this challenge, researchers use event log sampling, where a smaller subset of traces is selected as a representative proxy of the original log. This enables more resource-intensive process mining techniques to produce meaningful results on these smaller logs.

However, sampling from an event log can introduce changes, such as alterations in the contained activities, number of variants, or their frequency distribution. These changes may impact the results of subsequent process analysis, raising concerns about the reliability of the findings. Different sampling techniques, like random sampling, temporal sampling, or stratified sampling, affect the log in diverse ways, leading to variations in the potential effects on the analysis results. The objective of this seminar thesis is to explore the sampling techniques used in process mining research and study their potential detrimental effects on applied process mining techniques.

Literature:

- [1] Knols, B., & van der Werf, J.M E. (2019). [Measuring the behavioral quality of log sampling](#). International Conference on Process Mining.
- [2] Kabierski, M., Nguyen, H. L., Grunske, L., & Weidlich, M. (2021, October). [Sampling what matters: relevance-guided sampling of event logs](#). International Conference on Process Mining.
- [3] Fani Sani, M., van Zelst, S. J., & van der Aalst, W. M. (2021). [The impact of biased sampling of event logs on the performance of process discovery](#). Computing 103, 1085-1104.

Topic 6: Object-Centric Conformance Checking

Most existing conformance checking techniques assume that a process can be defined by a single case notion, i.e., all actions are related to exactly one object define a process instance. However, a process might only be conforming if multiple objects fulfill desired criteria. For example, consider a delivery process where multiple customers are supplied in multiple shipments with multiple items and the process is only conforming if all customers receive the prescribed quantity of all items. To account for multiple objects in a process instead of using only one case identifier, object-centric process mining with conformance checking applications has been proposed.

The goal of this seminar thesis is to identify techniques that conduct the conformance check using object-centric event-data. The applicant should identify 1) what are the inputs and outputs of object-centric process mining (including event log and process model formats), 2) what kind of technique is applied (e.g., alignments, token-replay), and 3) why the authors of each paper argue that their technique is better than previous research. Since this is a relatively new research topic, existing literature is expected to be sparse. Therefore, the participant should extensively discuss challenges and limitations of conformance checking in the context of object-centric process mining.

Literature:

- [1] Adams, J.N., & van der Aalst, W.M.P (2021). [Precision and fitness in object-centric process mining](#). International Conference on Process Mining.
- [2] Liss, L., Adams, J.N., & van der Aalst, W.M.P (2023). [Object-Centric Alignments](#). arXiv preprint arXiv:2305.05113.

Topic 7: Ex-Post-Analysis of Conformance Checking Results

Conformance checking is one of the main tasks of process mining. It comprises techniques for checking the relation between a designed process model and the real-life behavior of a process, with the goal of identifying and analyzing deviations between them. State-of-the-art conformance checking approaches have become computationally very efficient and robust to many kinds of process behavior. However, as the conformance checking itself cannot be considered as a means to an end, further analyses can prove useful in order to explain why deviations from the designed process model occur. Users will be enabled to improve the process conformance by that.

The goal of this seminar thesis is to identify approaches for the ex-post analysis of conformance checking results. The applicant should identify 1) which level of granularity in conformance measures is used (i.e., fitness, individual deviations, or deviation patterns like swapped activities), 2) for what purpose the conformance is analyzed, and 3) which methods are used to explain conformance differences. The review should be conducted wholistically, and the participant should contact the supervisor in case there is too much relevant literature (in which case the scope will be redefined). In a second step, the identified approaches should be analyzed with regard to a research gap or patterns in usage of conformance checking results.

Literature:

- [1] De Leoni, M., van der Aalst, W.M.P., & Dees, M. (2016). [A general process mining framework for correlating, predicting and clustering dynamic behavior based on event logs](#). Information Systems 56, 235-257.
- [2] Accorsi, R., & Stocker, T. (2012). [On the exploitation of process mining for security audits: the conformance checking case](#). Proceedings of the 27th Annual ACM Symposium on Applied Computing, 1709–1716.

Topic 8: The Application of Process Mining to Non-Process-Related Data

The goal of process mining is to analyze and improve business processes, as they occur in organizations. However, from a purely data-centric perspective, process mining can also be applied to event logs that do not refer to the execution of a business process, but to other collections of activities that jointly realize a certain goal. For example, process mining has been applied to analyze user interactions with a graphical user interface, the success of students in an online learning environment, or even the movement of tectonic plates.

The goal of this seminar thesis is to provide an overview of the application of process mining in non-process-related settings. Therefore, the participant should identify recent papers that apply process mining to study phenomena that are not business processes in a traditional sense. Those papers should be analyzed regarding (1) the phenomenon in focus, (2) the data they require, and (3) the process mining techniques they apply.

Literature:

- [1] Malinova, M., Gross, S., & Mendling, J. (2019). [Researching Information Systems Methods using Method Mining-A Case Study on Process Improvement Methods](#). International Conference on Information Systems.
- [2] Esnaashari, S., Gardner, L. A., & Rehm, M. (2022). [Exploring the Cyclical Nature of Self-Regulation in Blended Learning: A Longitudinal Study](#). International Conference on Information Systems.
- [3] Abb, L., Bormann, C., Aa, H. van der, & Rehse, J.-R. (2022). [Trace Clustering for User Behavior Mining](#). European Conference on Information Systems.

Topic 9: Reinforcement Learning in Business Process Management

Reinforcement Learning constitutes one of the three major categories of Machine Learning techniques. It has been successfully applied in various domains, for example to achieve superhuman performance in games like Go or to fine-tune conversational Language Models like ChatGPT. In Business Process Management and Process Mining, reinforcement learning could also be used for various purposes, such as autoregressive process prediction, fine-tuning process discovery results, and supporting process improvement efforts. However, its potential applications have only sporadically been investigated, and most machine learning systems in the field rely on supervised or unsupervised techniques instead.

The objective of this seminar thesis is to review existing research on reinforcement learning in the BPM domain. This involves identifying for each publication the problem that is to be solved, the general reinforcement learning approach applied (e.g., model-based or model-free), and ideally also why reinforcement learning is a suitable method to find a solution to the problem.

Literature:

- [1] Huang, Z., van der Aalst, W.M.P., Lu, X., & Duan, H. (2011). [Reinforcement learning based resource allocation in business process management](#). Data & Knowledge Engineering 70(1), 127-145.
- [2] Kurz, A.F., Kampik, T., Pufahl, L., & Weber, I. (2023). [Reinforcement Learning-Supported AB Testing of Business Process Improvements: An Industry Perspective](#). International Conference on Business Process Modeling, Development and Support.
- [3] Chiorrini, A., Diamantini, C., Mircoli, A., & Potena, D. (2020). [A Preliminary Study on the Application of Reinforcement Learning for Predictive Process Monitoring](#). ICPM Workshops.