



Bachelorarbeit / Masterarbeit

Event representations in relational databases

Process-oriented data mining (process mining) uses algorithms and data (in the form of event logs) to construct models that aim to provide insights into organisational processes. The quality of the event log data (both form and content) presented to the modelling algorithms is critical to the success of the process mining exercise.

However, these event logs are frequently extracted from databases, as typical business information systems do not provide process data in the necessary format. Event data recorded by information systems is often stored in relational tables with records connected through key relationships. References to cases and events are only implicitly available and have to be identified, extracted from the raw data and converted into flat event logs.

Event data can be encoded in multiple fashions in relational data and any database column with a *datetime* datatype is a pointer to event data. However, a general framework for constructing events from relational data is still missing. This thesis focuses on creating an exhaustive list of possible event representations in relational databases so that future event log extraction can be supported or even automated.

The thesis can be written in either English or German language.

Empfohlene Einstiegsliteratur:

- van der Aalst W. et al. (2012) Process Mining Manifesto. In: Daniel F., Barkaoui K., Dustdar S. (eds) Business Process Management Workshops. BPM 2011. Lecture Notes in Business Information Processing, vol 99. Springer, Berlin, Heidelberg
- S. Suriadi, R. Andrews, A.H.M. ter Hofstede, M.T. Wynn (2017) Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs. Information Systems (64, pp. 132-150)

Betreuer: van Dun, Christopher, M.Sc.