Unraveling the Fabric of Intertwined Processes: How Object-Centric Process Mining is Changing the Way We Improve Operational Processes (Invited Paper)

Wil M.P. van der Aalst[1,2]

[1] Process and Data Science (PADS), RWTH Aachen University, Germany
wvdaalst@pads.rwth-aachen.de
[2] Celonis, München, Germany

Abstract. While traditional process mining is powerful and effective, it also has its limitations. The extraction and transformation of data can be both time-consuming and repetitive, especially when changing viewpoints. Traditional event logs do not capture interactions between objects (such as sales orders, items, shipments, invoices, etc.), making it challenging to analyze these interactions using conventional methods. Yet, these interactions are often the source of problems. Additionally, fitting data into classical case-based event logs leads to the well-known convergence and divergence issues. Object-Centric Process Mining (OCPM) and the newly introduced OCEL 2.0 standard are designed to address these challenges. This extended abstract, derived from a keynote presented at the Workshop on Collaboration Mining for Distributed Systems (COMINDS 2023), offers insights into these recent advancements.

Keywords: Object-centric process mining · Concurrency · Object-centric event data · Rainbow Spaghetti · OCEL 2.0.

1 A Brief History of Process Mining

Process mining started in the late 1990-ties and initially focused on control-flow discovery, e.g., creating Directly-Follows Graphs (DFGs) or Petri nets describing the sequences of events seen in the databases of an organization [1].

In the period 2004–2011, the "golden age of process mining", many elements were added (see Fig. 1). The scope was extended to include conformance checking, decision mining, predictive analytics, process recommendations, and the automatic creation of simulation models. Also, the temporal, data, and resource perspectives were added.

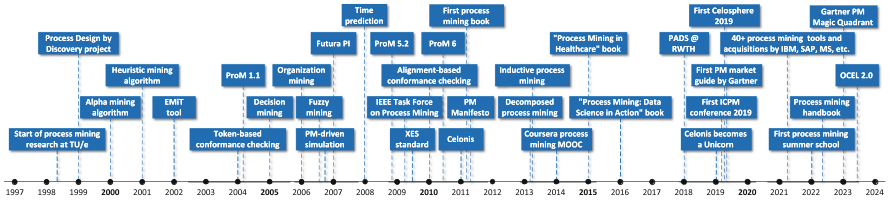

Fig. 1. A timeline of process mining showing important milestones.

In 2011, the "Process Mining Manifesto" was published [5] and the first version of the XES standard (cf. xes-standard.org) was released. Also, many process mining companies were founded (e.g., Celonis) leading to the over 40 commercial process mining products available today (see processmining.org).

In recent years, we have witnessed the uptake of process mining with thousands of organizations using it to improve their processes [3]. Although the process mining discipline has matured (see for example successful annual conferences such as ICPM and Celosphere) and there are many success stories, we can also see the limitations of the initial assumption that there is a single case notion and each event refers to precisely one case. This triggered the development of *Object-Centric Process Mining* (OCPM) [2] discussed next.

2 A New Standard for Object-Centric Event Data

Object-Centric Event Logs (OCELs) form the basis for Object-Centric Process Mining (OCPM). OCEL 1.0 was first released in 2020 and triggered the development of a range of OCPM techniques. OCEL 2.0 forms the new, more expressive standard, allowing for more extensive process analyses while remaining in an easily exchangeable format.

In contrast to the first OCEL standard, OCEL 2.0 can depict changes in objects, provide information on object relationships, and qualify these relationships to other objects or specific events (see Fig. 2). Compared to XES, it is more expressive, less complicated, and better readable. OCEL 2.0 offers three exchange formats: a relational database (SQLite), XML, and JSON format. See [4] for the specification and [6] for datasets and tool support.

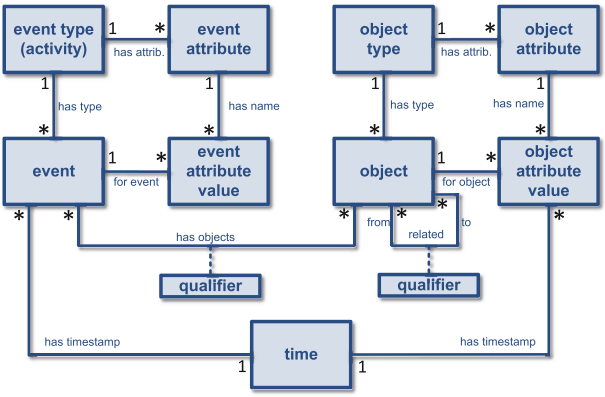

Fig. 2. The OCEL 2.0 Meta Model [2, 4, 6].

3 Advantages of Using Object-Centric Process Mining

For a basic introduction to Object-Centric Process Mining (OCPM) and an overview of literature, we refer to [2]. Here we limit ourselves to discussing the advantages of using OCEL 2.0 and related OCPM techniques.

Real-life processes can be seen as "Rainbow Spaghetti" where each strand of Spaghetti (called "Spaghetto" in Italian) corresponds to the lifecycle of an object and the color refers to the type. Often, it is not sufficient to limit the view to one object type, because the root causes of problems often involve multiple objects of different types. Therefore, classical process mining involves mapping multiple object types on a single case notion.

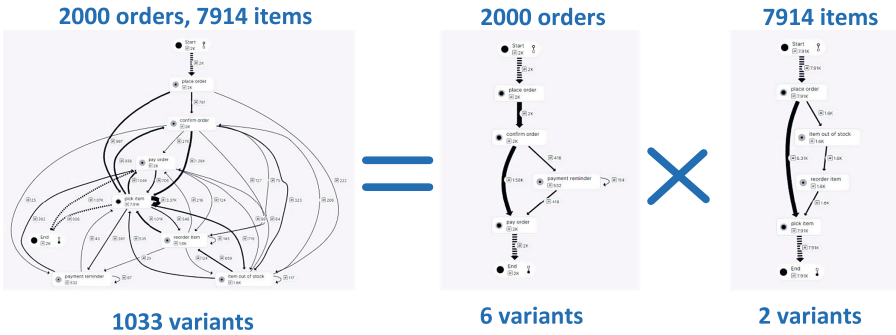

Fig. 3. Example showing the drawbacks of mapping different types of objects onto a single case notion. Although the handling of orders and items in isolation is highly structured, the combined model using a single-case notion is complex and misleading.

Figure 3 shows three Directly-Follows Graphs (DFGs). The two DFGs on the right-hand side show that the handling of orders and items in isolation is simple. In a data set with 2000 orders and 7914 items, there are just 6 respectively 2 variants. However, if both are combined, we get the DFG on the left with 1033 variants. The combined DFG uses orders as a case notion, but also includes the activities that happened for the items in the order. Orders get delayed by items being out of stock; therefore, we need to consider also these activities. However, the resulting DFG has convergence and divergence problems [2]. Statistics related to frequencies and times are distorted. Moreover, the DFG also shows connections such as "pick item" followed by "reorder item" and "item out of stock". These connections are not causal and are created by squeezing two types of objects into one case notion. Using OCEL 2.0 and OCPM, one can avoid such problems. Object-centric event data reflect reality without distortions and the discovered multi-object DFGs do not suffer from convergence and divergence problems.

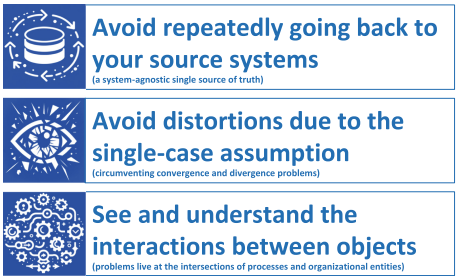

Fig. 4. Overview of the three main reasons to use OCPM.

Figure 4 summarizes the three main advantages of OCPM. Using OCEL 2.0, one can store object and event data in a system-agnostic manner. Therefore, extraction does not need to be repeated each time the view or case notion changes. Figure 3 showed the problems when squeezing reality into a model using a single-case notion. This generally leads to misleading diagnostics that can only be understood correctly by experts who transformed the source data. Finally, using OCPM it is possible to see and understand interactions between objects of different types. Performance and compliance problems cannot be understood by looking at objects in isolation.

Acknowledgments. Supported by the Alexander von Humboldt (AvH) Stiftung and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy, EXC Internet of Production, 390621612. Thanks to the PADS team for driving the development of the OCEL 2.0 standard [4] and OCPM techniques.

References

1. van der Aalst, W.: Process Mining: Data Science in Action. Springer, Heidelberg (2016). https://link.springer.com/book/10.1007/978-3-662-49851-4
2. van der Aalst, W.: Object-centric process mining: unraveling the fabric of real processes. Math. **11**(12), 2691 (2023)
3. van der Aalst, W., Carmona, J.: Process Mining Handbook. Lecture Notes in Business Information Processing, vol. 448. Springer, Cham (2022). https://link.springer.com/book/10.1007/978-3-031-08848-3
4. Berti, A., et al.: OCEL (Object-Centric Event Log) 2.0 Specification (2023). https://www.ocel-standard.org/2.0/ocel20_specification.pdf
5. IEEE Task Force on Process Mining: Process Mining Manifesto (2011). https://www.tf-pm.org/resources/manifesto
6. Process and Data Science Group: Object-Centric Event Log (OCEL 2.0) Website Providing Data Sets, Software, Etc. (2023) . https://www.ocel-standard.org/