

# Object-centric Process Mining on Collection Objects

Jost Götte<sup>1,\*</sup>

<sup>1</sup>*Kühne-Logistics-University, Hamburg, Germany*

## Abstract

Process Mining, a tool to analyze event data, traditionally extracts this data in form of flat event logs. This results in limitations like divergence and convergence. Object-centric process mining (OCPM) overcomes this by modeling objects and their relations explicitly. However, objects are often treated as strictly singular entities. Instead objects could be made up of multiple other objects. We refer to them as collection objects. This project's goal is to establish fundamental understanding about collection objects (RQ1), develop novel OCPM techniques to solve inherent challenges of processes with collection objects (RQ2) and apply these novel techniques to real world supply chain management (SCM) settings, to evaluate the feasibility of these techniques and gather new insights for the SCM domain (RQ3).

## Keywords

Object-centric Process Mining, Supply Chain Management, Collection objects, Temporal relations

## 1. Introduction

Process Mining offers methods and techniques to extract, analyze and improve business processes. It finds application in areas like healthcare, finance or supply chain management (SCM) [1]. Traditionally, Process Mining extracts event data from its data sources in the format of flat event logs. Objects and their relations are flattened to represent the perspective of one object type, which becomes the case of the process.

As shown in [2] this leads to issues like divergence and convergence, in cases where objects interact with multiple other objects of the same object type at once. Object-centric Process Mining (OCPM) addresses these issues by substituting the case notion for explicit object relations. OCPM techniques are able to infer valuable information about the process via these object relations.

## 2. Collection objects

One class of objects that benefits from this paradigm shift are collection objects. These are objects that consist of a group of objects. A typical order for example may consist of multiple goods. Typically, process mining refers to these goods only implicitly by associating the order with a quantity and a type attribute.

Intuitively, this is sensible in instances where the process on the level of these goods is not relevant for the Process Mining analysis. However, if we consider larger orders where goods are aggregated into (line-)items and these items exhibit deviating behavior, it becomes relevant to incorporate these items in the process. In this example, the order object is related to the item objects it consists of. As a result, the process can be analyzed on multiple levels, the order level and the item level. This multi-level event execution, where collection objects and their elements occasionally share events and occasionally deviate in their behavior, is the first special challenge of working with collection objects. Similar multi-level concerns can be applied in hierarchical process mining [3]. The novelty for collection objects is that the multi-level view extends to the objects of the process as well, the collection objects.

---

*ICPM Doctoral Consortium and Demo Track 2025, October 20-24, 2025, Montevideo, Uruguay*

\*Corresponding author.

✉ [jost.goette@klu.org](mailto:jost.goette@klu.org) (J. Götte)

🌐 <https://github.com/Jostafarr> (J. Götte)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

The second challenge does not come from the structure of the collection objects itself but from their characteristic behavior. Regardless whether collection objects are modeled explicitly or implicitly, they can be subject to several operations. Their quantity might be increased or decreased, or they might exchange elements with other collection objects. These operations are easier to implement in process mining techniques when collection objects are modeled implicitly through activities that increase/decrease their quantity accordingly. In instances, where collection objects are modeled explicitly, an increase or decrease means that one object is now related to a new item or a relation ceases to exist. Such dynamic relations are a new concept to OCPM that models relations without regard to time-bound constraints.

This gap increases in severity when considering more complex operations such as replacements, fragmentation, or consolidation. Here, one or more collection objects cease to exist and one or more collection objects take their place. For instance, fragmentation might be used to split an order into smaller orders with different delivery dates. For any of these operations the involved collection objects are temporally related as part of a predecessor-successor relation. When collection objects are implemented explicitly this means that all predecessor elements are now related to the corresponding successor elements, highlighting the increased complexity of the dynamic relation. Temporal relations between objects are still relatively uncovered by OCPM research but not unknown. A temporal object type model (TOTEM) [4] reveals high-level timing dependencies between object types, although they do not address intra-type hierarchical temporal dynamics such as those between collection objects.

Additionally, this second group of operations highlights another challenge, that is their mismatch of object life-cycles. When one object is replaced by another the life-cycle of the first object ends and the other starts, however the linked elements persist. Counterintuitively, all of this is part of one process instance that has to be traced over multiple object life-cycles.

### 3. Research Questions

Considering the challenges positioned above, we ask how they can be answered in OCPM. We seek to answer the following three research questions:

1. What are collection objects in the context of OCPM?
2. How can OCPM overcome the challenges of mining collection objects?
3. What novel insights can OCPM for collection objects generate when applying it to real world supply chain data?

These questions do not necessarily need to be answered in the exact order they are presented, although they rely on each other to a certain degree. Rather, we envisage an iterative approach, hoping that the insights we gain from applying OCPM techniques to collection object data in SCM may lead to new questions and gaps to address and to create knowledge about which again leads to improved or novel OCPM techniques. Ideally, SCM acts as a demonstration area and while we also use it to motivate the topic of collection object, our goal is to develop concepts that are applicable outside this domain.

#### 3.1. RQ1

Collection objects have not been addressed directly by current (object-centric) process mining techniques. The goal of RQ1 is to gain fundamental understanding of collection objects, their behavior and corresponding challenges.

Therefore, it is necessary to survey the current state of supply chain and process mining literature to expose gaps and opportunities for implementing support for collection objects from both perspectives. The supply chain side ensures relevance and compatibility of collection objects for the field, while the process mining side helps to uncover technical and conceptual limitations of current Process Mining methods.

Furthermore, collection objects need to be formally defined and characterized in order to create universal understanding and as a basis to enable research on them. This can be achieved using mathematical formalizations that build on the existing definitions of OCPM like the OCEL 2.0 specification [5], to ensure that collection objects are directly immersed into their object-centric context, including an potential extension of the object-centric metamodel.

### **3.2. RQ2**

Once the preliminary understanding about the concept of collection objects is established, RQ2 seeks to address the identified gaps and challenges (see section 2. This includes the integration and adaptation of existing OCPM techniques, as well as deriving novel OCPM techniques. The resulting contribution will lie in the technical artifacts that are created along the way.

Under RQ2 all four process mining tasks could be investigated. However, we will limit ourselves to process discovery and process improvement. This focus fits well to the application domain SCM. It includes finding suitable process representations for modeling and logging processes with collection objects. For instance, one could compare object-centric visualizations with event-knowledge graphs [6].

With regard to the benefits of this research for supply chain management, the process improvement tasks of Process Mining is especially relevant. Establishing (object-centric) process performance indicators [7] might lead to the development of new KPIs to improve understanding of supply chain processes.

Hence, methodologies to address RQ2 include inductive approaches, based on the supply chain's domain perspective as well as the deductive approach of design science. Either way, the inherent challenges of collection objects need to be addressed in any resulting OCPM technique. Ideally, the evaluation of the technical artifacts produced under RQ2 would be based on real world data of SCM cases.

### **3.3. RQ3**

Studying this kind of application of the in RQ2 developed OCPM techniques on supply chain data and their implications, is the goal of RQ3. This would not only validate the applicability of these approaches but could also generate novel insights for the supply chain domain itself. This might include the definition of new KPIs or policies to order stock in collection object heavy processes. Current approaches that consider for example partial or split approaches do not consider larger numbers of splits. Usually they stick to two partial deliveries [?] Case studies as well as expert interviews with suitable industry partners can test the practical feasibility and validate our developed concepts. However, at this point in time, we do not have any agreements for such collaborations, yet.

## **4. Initial Steps**

In an initial effort, a pilot project has been conducted that investigates the usage of collection objects for process performance analysis on supply orders with partial deliveries. The goal of the pilot was to make a step towards RQ1 and RQ2 within the set limitations of the use case. Here, it was assumed that collection objects may only be split in the form of order fragmentation that leads to partial deliveries and do not exhibit other behavior like consolidation. As these order fragmentations produce the object life-cycle mismatches, that span a directed acyclic graph, breadth-first search has been applied so that related orders could be traced from the starting order. At the same time, we have proposed a formalization for collection objects and the fragmentation operation. The approach has been evaluated using synthetic supply data that imitates a real-world process.

The synthetic data, used in the project above, was created by a tool that we designed to enable OCPM research on collection objects but also supply data in general. It simulates a virtual warehouse using the continuous review stock policy [8], that reorders stock, whenever it drops below a certain threshold. The simulated orders are then gathered to create an OCEL, that is the output-artifact of the tool. To

simulate the special behavior of collection objects, the tool offers to customize to which degree orders can be split. In the future, it could be adapted to offer other kinds of collection object behavior.

This pilot study was as a bottom up approach that taught us the relevancy of research on collection objects on a smaller context, more specific problem definition. Now, the challenge for us lies in tackling the topic of collection objects from a top down perspective to keep contributions we make generalizable to a greater context than the supply chain context, we set out from.

## 5. Conclusion

This research addresses the different facets of collection objects in the realm of OCPM. Its goal is to create understanding for the current challenges and gaps (RQ1), develop novel and adapt existing OCPM techniques to address these challenges (RQ2) and then apply them to the domain of supply chain management to gather new insights about real world supply chain obstacles. While some initial efforts have been done, the majority of research opportunities are yet untouched.

## Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

## References

- [1] F. Oldenburg, K. Hoberg, H. Leopold, Process mining in supply chain management: state-of-the-art, use cases and research outlook, *International Journal of Production Research* 63 (2025) 2889–2904. URL: <https://doi.org/10.1080/00207543.2024.2412285>. doi:10.1080/00207543.2024.2412285, publisher: Taylor & Francis \_eprint: <https://doi.org/10.1080/00207543.2024.2412285>.
- [2] W. M. P. Van Der Aalst, Object-Centric Process Mining: Dealing with Divergence and Convergence in Event Data, in: P. C. Ölveczky, G. Salaün (Eds.), *Software Engineering and Formal Methods*, volume 11724, Springer International Publishing, Cham, 2019, pp. 3–25. URL: [http://link.springer.com/10.1007/978-3-030-30446-1\\_1](http://link.springer.com/10.1007/978-3-030-30446-1_1). doi:10.1007/978-3-030-30446-1\_1, series Title: Lecture Notes in Computer Science.
- [3] S. J. Leemans, K. Goel, S. J. van Zelst, Using Multi-Level Information in Hierarchical Process Mining: Balancing Behavioural Quality and Model Complexity, in: *2020 2nd International Conference on Process Mining (ICPM)*, 2020, pp. 137–144. URL: <https://ieeexplore.ieee.org/abstract/document/9230000>. doi:10.1109/ICPM49681.2020.00029.
- [4] L. Liss, J. N. Adams, W. M. P. Van Der Aalst, TOTeM: Temporal Object Type Model for Object-Centric Process Mining, in: A. Marrella, M. Resinas, M. Jans, M. Rosemann (Eds.), *Business Process Management Forum*, volume 526, Springer Nature Switzerland, Cham, 2024, pp. 107–123. URL: [https://link.springer.com/10.1007/978-3-031-70418-5\\_7](https://link.springer.com/10.1007/978-3-031-70418-5_7). doi:10.1007/978-3-031-70418-5\_7, series Title: Lecture Notes in Business Information Processing.
- [5] A. Berti, I. Koren, J. N. Adams, G. Park, B. Knopp, N. Graves, M. Rafiei, L. Liß, L. T. G. Unterberg, Y. Zhang, C. Schwanen, M. Pegoraro, OCEL (Object-Centric Event Log) 2.0 Specification (????).
- [6] D. Fahland, Process Mining over Multiple Behavioral Dimensions with Event Knowledge Graphs, in: W. M. P. Van Der Aalst, J. Carmona (Eds.), *Process Mining Handbook*, volume 448, Springer International Publishing, Cham, 2022, pp. 274–319. URL: [https://link.springer.com/10.1007/978-3-031-08848-3\\_9](https://link.springer.com/10.1007/978-3-031-08848-3_9). doi:10.1007/978-3-031-08848-3\_9, series Title: Lecture Notes in Business Information Processing.
- [7] B. Estrada-Torres, A. del Río-Ortega, M. Resinas, Defining Process Performance Measures in an Object-Centric Context, in: C. Cabanillas, N. F. Garmann-Johnsen, A. Koschmider (Eds.), *Business Process Management Workshops*, volume 460, Springer International Publishing, Cham, 2023, pp. 210–222. URL: [https://link.springer.com/10.1007/978-3-031-25383-6\\_16](https://link.springer.com/10.1007/978-3-031-25383-6_16). doi:10.1007/978-3-031-25383-6\_16, series Title: Lecture Notes in Business Information Processing.

- [8] S. Axsäter, Inventory control, number 90 in International series in operations research & management science, 2nd ed ed., Springer, New York, 2006.