Providing Semantics to Object-Centric Event Logs (OCEL) from Process Mining using BFO

Evellin Cardoso[1]

[1]Federal University of Goias, Goias, Brazil

Abstract

Business Process Management (BPM) is a discipline encompassing a wide array of methods, techniques, and tools derived from both Information Technology and Management Sciences to manage business processes within organizations. This paper addresses a critical BPM challenge, specifically concerning the semantic enrichment of process data. We propose a novel framework designed to augment Object-Centric Event Logs (OCEL), a recent advancement in Process Mining, with the robust ontological foundation provided by the Basic Formal Ontology (BFO). By integrating OCEL data with BFO, our framework seeks to enrich process mining artifacts with deeper semantic meaning, thereby enabling more sophisticated analysis, more comprehensive understanding of business processes, also opening the possibility to obtain richer insights from Process Mining algorithms.

Keywords

Ontology, Basic Foundational Ontology (BFO), Process Mining, Data, Knowledge-Augmented Business Process Management (BPM)

1. Introduction

The Business Process Management (BPM) discipline provides a structured approach to understanding, optimizing, and controlling the business processes that operate in enterprises [1]. Central to BPM is the concept of a business process, which can be understood as a coordinated set of activities performed to handle a specific case or process instance [1]. The BPM lifecycle typically involves several phases: (i) the (re)design phase, in which process models are elaborated; (ii) the configuration phase, where these models are implemented within a Business Process Management Systems (BPMS), and (iii) the execution phase, during which numerous process instances are created, with sequences of activities reflecting the evolution of each instance. Traditionally, process analysis has heavily relied on classical event logs, which capture activities identified by a single case ID.

The field of Process Mining, which systematically utilizes event data to improve business processes [3] witnessed the emergence of Object-Centric Process Mining (OCPM) techniques. OCPM introduces Object-Centric Event Logs (OCEL) [4], a paradigm shift in how event data is organized. Unlike classical event logs that are case-centric, OCELs are structured in relation to the various data objects involved in a business process, rather than solely focusing on a single case notion. While OCEL offers a more granular and interconnected perspective on data, it still predominantly operates within the procedural sphere, focusing on timestamped activities and how they update data object attributes, without explicitly incorporating the broader enterprise knowledge structure.

To tackle this gap between knowledge and data, this paper addresses the challenge of augmenting the event data stored in OCEL logs with process and domain knowledge. To achieve this, we propose a novel framework that integrates OCEL data with ontologies, specifically leveraging the Basic Formal Ontology (BFO) [5]. Ontologies, as a cornerstone of Semantic Web technologies, provide a common vocabulary for representing knowledge and information across heterogeneous resources and applications, thereby promoting data integration and interoperability. BFO, an ISO standard since 2002, is chosen for its

Proceedings of the 18th Seminar on Ontology Research in Brazil (ONTOBRAS 2025) and 9th Doctoral and Masters Consortium on Ontologies (WTDO 2025), São José dos Campos (SP), Brazil, September 29 – October 02, 2025.
*Corresponding author.
evellin@ufg.br,evellinc@gmail.com (E. Cardoso)
0000-0001-6242-662X (E. Cardoso)

realist stance to capture general features of reality. By bridging OCEL with BFO, our framework aims to provide a semantically rich representation of business processes, enabling a deeper level of analysis and insight that transcends the limitations of purely procedural data. This integration thereby paves the way for more intelligent and robust process management solutions.

This paper is structured as follows. Section 2.1 introduces the BPM discipline and the Basic Foundational Ontology (BFO). Section 3 presents the OntoOCEL framework that augments data from OCEL logs with BFO concepts. Section 5 concludes the paper, outlining directions for future work.

2. Baseline

2.1. Business Process Management (BPM) and Process Mining

Business Process Management (BPM) is the discipline that includes methods, techniques and tools from Information Technology and Management Sciences to support the design, analysis, configuration, execution and monitoring of business processes [1].

The notion of business process is central to BPM. A *business process* consists of a set of activities that are executed to handle a *case* (a process instance) [1]. To manage business processes, BPM conducts its efforts along different phases of the *BPM lifecycle* [1]. In the (re)design phase, a process model is elaborated. This model is subsequently configured/implemented in a business process management system (BPMS) within the configuration phase. Within the execution phase, processes instances are created multiple times, with sequences of activities corresponding to the evolution of each case.

Taking advantage of this event data and process models, *Process Mining* is the discipline that systematically uses event data to improve business processes [3]. Traditional process mining techniques rely on classical event-logs which are based on the assumption of a single case notion, with events referring to exactly this case. More recently, Object-Centric Process Mining (OCPM) introduces Object Centric Event Logs (OCEL) which are organized in relation to the data objects present in the business, rather than in relation to the cases as a sequence of events. Although this novel way of organizing events in OCEL reflects a more seamless way to look into the data, which is closer to the process as it actually is, the data analyzed by OCPM algorithms still heavily remain in the processual sphere of the business.

The fact that only procedural knowledge is tracked in BPMS system has been already identified as one of the major unresolved issues in BPM [2]. In this work, authors highlight that event logs lack domain-specific and commonsense knowledge, which make them often suffering from noise and incompleteness, thus impacting the outcomes and insights provided by traditional process mining algorithms. To tackle this problem, we augment the event data from OCEL logs with process and domain knowledge in this paper.

2.2. Ontologies and Semantic Technologies

Recently, Semantic Web technologies started gaining increasing attention for their ability to promote heterogeneous data integration and interoperability. Ontologies provide a common vocabulary for representing knowledge and information across heterogeneous resources and applications.

To support such interoperable environment, different types of ontologies exist [5]. *Upper-level ontologies* provide a highly general vocabulary of categories and relations regardless of domain, while *domain ontologies* covers a basic set of universal categories from particular scientific domains.

There exists a number of upper-level ontologies already in place, such as DOLCE, UFO, GFO, BFO, etc [5]. The Basic Foundational Ontology (BFO) is chosen in this paper to promote semantic interoperability because it has been introduced as an ISO standard since 2002, including BFO's ISO 21838-2 specification being axiomatized in First-Order Logic, OWL 2, and CLIF [6].

BFO top-level ontology has been designed with a realist instance in mind. Its main goal is to use ontologies to represent the knowledge acquired as a result of scientific efforts. Being realist means that such knowledge captures general features of reality as general theories (i.e. generalizations and laws of science), rather than particular facts. To make such distinction, BFO presents some fundamental categories as described below:

Universal, **Particular** and **Defined Classes.** Universals are mind-independent entities that can be repeatedly instantiated across time and space, with an indefinite number of particulars, while particulars are individual entities, restricted to specific places. They instantiate universals, but cannot be instantiated. Defined classes are general terms used in science to refer to particular individuals in reality (e.g., medical doctor, dog) that have no corresponding universals [5]. In this way, medical doctor would simple be an instance of person (an Universal) that bear role of medical doctor.

Relations: to relate these entities, BFO introduces three basic relations: (i) universal-universal, (ii) universal-particular and (iii) particular-particular. While universal-universal relations connect subtypes to parent types (*IS A* relation), universal-particular relations relate the instances (particulars) to the universals in which they fall (*instance of*).

BFO is structured in terms of two disjoint hierarchies of universals, depending on how particulars relate to time. Continuants and occurrents are the roots of each branch:

Continuant and **Occurrent:** Continuants endure through time, retaining their identity, fully existing at any time they exist. Examples include a house, an apple, color of an orange. In contrast, occurrents unfold over time, being composed of temporal parts. Examples include a talk, a race, the history of Brazil, a period of time in which the sun rises, people attending a meeting.

Process and **Temporal instant:** A process is an occurrent composed by some temporal proper parts, a time interval $I = [t_i, ..., t_f]$ in which a *material entity* is as *participant*, while a temporal instant is a zero-dimensional temporal region with no temporal parts.

In scientific research, it is relevant to categorize different types of information entities that carry scientific knowledge. For this reason, the Information Artifact Ontology (IAO) is an ontology created using BFO as a basis, capturing numerous information entities existent in scientific research (e.g., protocols, documents, experimental logs, databases, published literature and so on).

Information Content Entity. In IAO, an information content entity is a *generically dependent continuant* that refers to (*is about*) some entity.

Document. A number of information entities that must be understood together as a whole.

As an upper ontology, BFO provides the foundation for over 350 domain ontology extensions in multiple domains [6]. In this paper, BFO is used for augmenting data with semantics. We refrain from including a definition for all the concepts here required in Section 3, referring the reader to [5, 6] for a comprehensive definition.

3. OntoOCEL Framework

This section describes the OntoOCEL framework to augment OCEL event data with process and domain knowledge:

3.1. Step 1: Understand How Process Knowledge Relates to the Ontology

The first step of our methodology consists of conducting a semantic analysis to understand how the process knowledge structure can be mapped to the to upper level ontology, regardless process instances.

In BPM literature, a business process is composed of a set of *activities* executed to handle a case (Section 2.1) (the control-flow perspective). Activities and processes are performed by *roles* that can be fulfilled by people or organization units (resource perspective), manipulating data and information (data perspective), happening during a period of time (time perspective) [1].

An assessment of BFO classes (Section 2.2), together with use cases on how to relate particulars to BFO universals and defined classes defined in [6] lead to the modeling decisions depicted in Figure 1. In short, BPM concepts are mapped to BFO *particulars* (represented as diamonds). These particulars are related to BFO *universals* and *defined classes* (represented as ovals) through an *instance of* relation. The relations among the particulars are defined as follows. Roles (organizational or person) *participate in* activities at a timestamp t. Activities *consume* or *produce* data objects.

Example. The aforementioned modeling decisions are schematically presented in Figure 1.

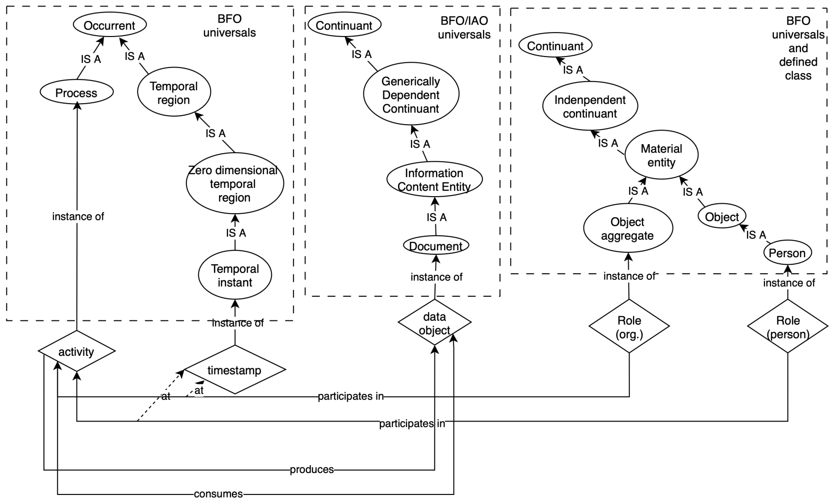

Figure 1: Process Knowledge Mapped to BFO Upper Ontology

3.2. Step 2: Understand the Semantics of Event Data in OCEL Logs

Now that we know how the process knowledge relates to the upper ontology, we can think about the domain knowledge. To do this, we initially look into the data stored in the event logs.

To facilitate interoperability among process mining tools, event logs are standardized. Figure 2(a) depicts the OCEL metamodel adapted from [4], while Figure 2(b) depicts a sample of synthetic, manually generated OCEL event log. As can be seen in Figure 2, OCEL event logs contain object types (and attributes), object instances, event types (activities) (and attributes) and event type instances.

With the OCEL event logs in hands, one can understand the data objects that we have available data, together with the events that manipulate such data. These data objects will be our domain entities in the next step.

Example. With the event logs in hands, it is possible to identify two data objects specified within the log ($Patient$ and $Prescription$) and one activity ($AdmitPatient$) □

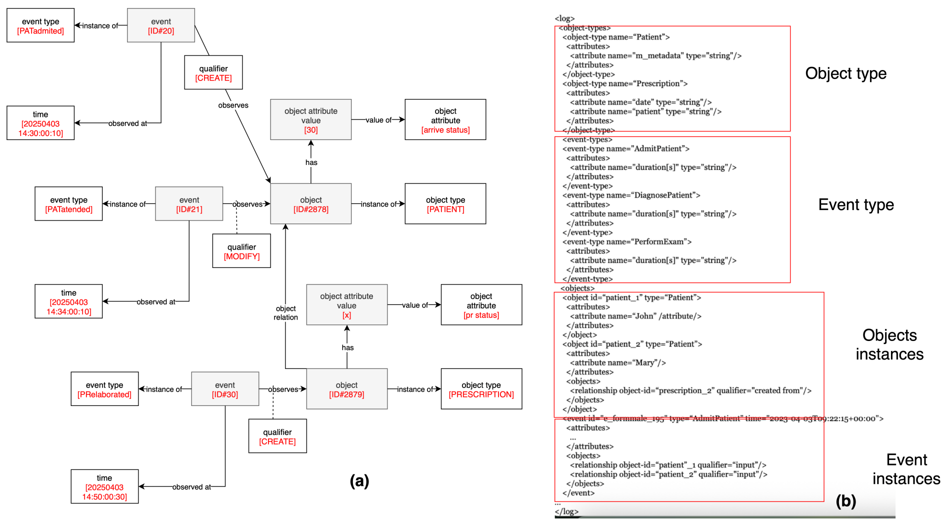

Figure 2: (a) OCEL metamodel adapted from [4] and (b) Sample of synthetic, manually generated OCEL event log

3.3. Step 3: Develop the Domain Ontology Based on Event Data

The third step consists of understanding the nature of knowledge stored in these data objects and how they fit with the knowledge structure of process knowledge (and indirectly to BFO). Activities are instantiated as subclass of BFO *process*, the data objects are instantiated as subclasses of BFO *information content entity*, roles are instantiated either as subclass of (defined class) *person* or as subclass of BFO *object aggregate*. The data objects also carry the values of the data object attributes. These are modeled as OWL Properties.

Example. Figure 3(a) depicts the hierarchy of BFO classes in OWL. First, the OCEL *object types* ($Patient$, $Prescription$ and $Treatment$) are instantiated as subclasses of *information content entity*. As one can see in the hierarchy, BFO provides different types of subclasses as carrier content (e.g. figures, documents, email, etc.). Here, the class $ElectronicRecord$ has been created as subclass of BFO *document* and pointed to class $Patient$ (subclass of BFO *object*), since a BFO information content entity *is about* some entity in reality. Relations among concepts are modeled as OWL Object Properties in Figure 3(b). Activities $AdmitPatient$, $PerformExam$ and $DiagnosePatient$ are created as subclass of BFO *process*. □

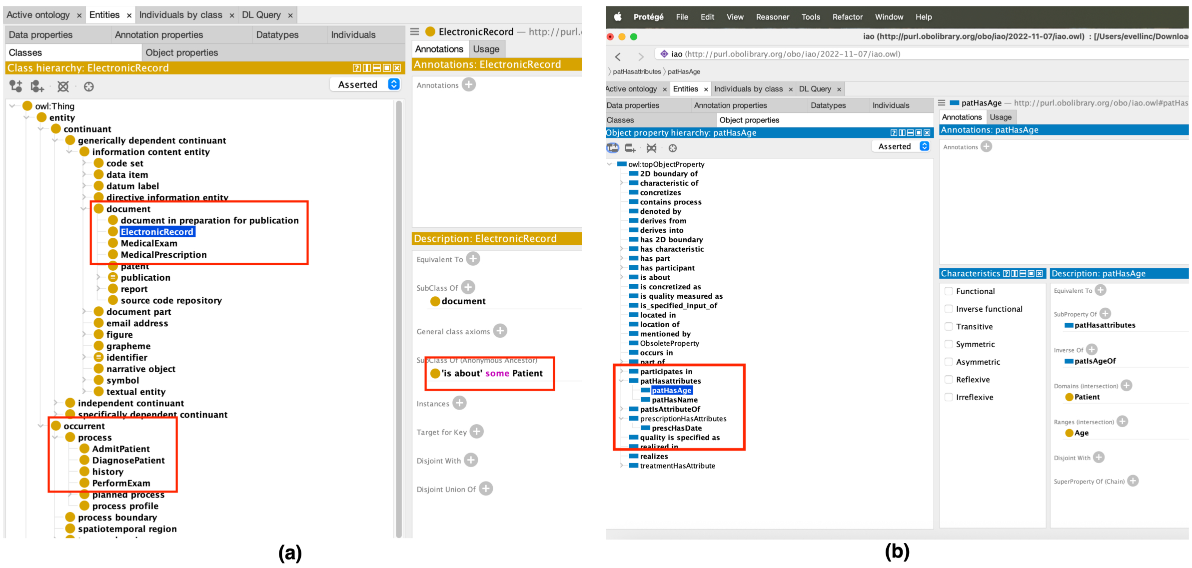

Figure 3: The Domain Ontology Linked to BFO Universals (1)

3.4. Step 4: Complement the Domain Ontology with Declarative Knowledge

In this step, the knowledge structure of the enterprise may be complemented with declarative knowledge. This knowledge should be related with the procedural knowledge modeled in the previous steps.

Example. Imagine an admission process that patients may have different types of diseases. The doctor would like to know about it, he wants to track it along time. In this case, we insert a $constitucionalGeneticDisease$ and $acquiredGeneticDisease$ as a disposition as depicted in [5]. Figure 4 depicts this modeling decision. □

3.5. Step 5: Integrate the Data with Ontologies

This step concerns the linkage between processual and domain data with the ontology. The Cellfie plugin [7] may be used for importing the OCEL event logs to the OWL ontology. Working with Cellfie requires the input data to be stored in an Excel spreadsheet, and the creation of *Transformation Rules* that maps the Excel data to the OWL axiom structure. The transformation rules are written in Manchester Syntax. Currently, an event log does not exist for the process, although the ontology has been already conceived in such a way to enable subsequent import of event logs. Further, domain specific data may

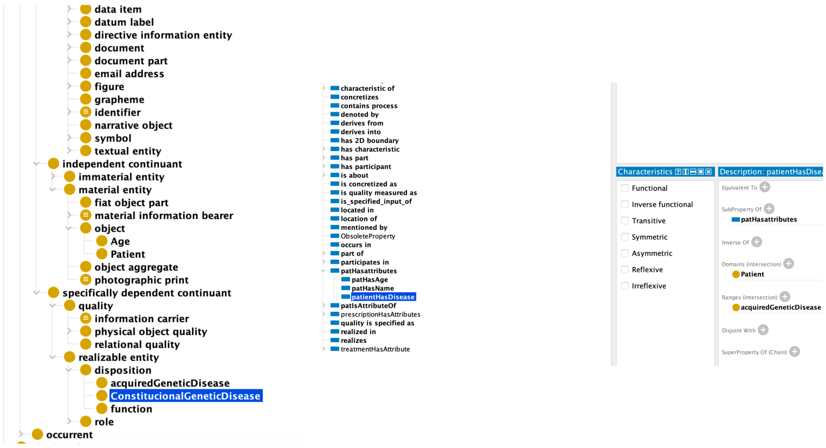

Figure 4: The Domain Ontology Linked to BFO Universals (2)

also be inserted into the ontology (e.g. data about the $constitucionalGeneticDisease$), extracted from relational databases, or other sources of information.

4. Related Work

The idea of providing semantics for BPM technologies has been embraced in two veins in literature. A first group of approaches use Ontologies and Semantic Web technologies combined with business process models [8, 9, 10, 11], augmenting the process logic with domain knowledge. These approaches are grounded on domain ontologies and OWL language, which does not embrace a semantic view supported by upper ontologies. The only exception is Pedrinaci et. al. [8] that considers upper ontologies, without deepening further on data.

In a second strain of research that augment event data with Semantic Web technologies [12, 13, 14, 15, 16], Ciccio et. al. [15] and Khayatbashi et. al. [16] are focused on temporal constraints, not on semantics. Eichele et. al. [12] indeed focus on semantics, but only consider domain ontologies (not upper) and classical event logs. Xiong et. al. [13] and Swevels et. al. [14] emphasize integration, extraction, and transformation, with weak semantic foundations and no foundational ontology grounding. Differently, this paper is explicitly dedicated to integrating OCEL data with the Basic Formal Ontology (BFO), proposing a framework that maps OCEL event structures to BFO categories to represent processes, participants, objects, and roles more richly. This approach enhance semantic meaning, interoperability, and support for knowledge-intensive process analysis, since it uses a well-leveraged ontology.

5. Conclusion

This paper has presented a comprehensive framework, OntoOCEL, designed to address the critical semantic gap in Process Mining by augmenting Object-Centric Event Logs (OCEL) with a rich ontological foundation derived from the Basic Formal Ontology (BFO). We have outlined a five-step methodology that facilitates this integration, moving from a semantic analysis of process knowledge in relation to upper-level ontologies to the development of a domain-specific ontology, its complementation with declarative knowledge, and finally, the seamless integration of event data with the constructed ontology.

While the current work lays a conceptual and methodological foundation, it is important to acknowledge that populating the ontology with event data is a limitation of our work that must addressed in future efforts. Future work will focus on the empirical validation of this framework through the actual import and analysis of real-world OCEL event logs, by elaborating the transformation rules, and exploring the potential for automated reasoning and intelligent decision support based on the enriched ontological models. This will ultimately lead to more intelligent, adaptive, and insightful BPMS systems.

References

[1] W. Aalst, van der, Business Process Management: A Comprehensive Survey, ISRN Software Engineering 2013 (2013).

[2] I. Beerepoot, C. Di Ciccio, H. A. Reijers, S. Rinderle-Ma, W. Bandara, Andrea, The Biggest Business Process Management Problems to Solve Before We Die, Computers in Industry 146 (2023) 103837. URL: https://www.sciencedirect.com/science/article/pii/S0166361522002330. doi:https://doi.org/10.1016/j.compind.2022.103837.

[3] W. van der Aalst, J. Carmona, Process Mining Handbook, Lecture Notes in Business Information Processing, Springer International Publishing, 2022. URL: https://books.google.com.br/books?id=0tMszwEACAAJ.

[4] D. Fahland, M. Montali, J. Lebherz, W. M. P. van der Aalst, M. van Asseldonk, P. Blank, L. Bosmans, M. Brenscheidt, C. di Ciccio, A. Delgado, D. Calegari, J. Peeperkorn, E. Verbeek, L. Vugs, M. T. Wynn, Towards a Simple and Extensible Standard for Object-Centric Event Data (OCED) – Core Model, Design Space, and Lessons Learned, 2024. URL: https://arxiv.org/abs/2410.14495. arXiv:2410.14495.

[5] R. Arp, B. Smith, A. D. Spear, Building Ontologies with Basic Formal Ontology, The MIT Press, 2015.

[6] J. N. Otte, J. Beverley, A. Ruttenberg, Basic Formal Ontology: Case Studies, Applied Ontology 17 (2022) 17–43. doi:10.3233/ao-220262.

[7] M. O'Connor, Cellfie plugin, https://github.com/protegeproject/cellfie-plugin/wiki/Grocery-Tutorial, 2025.

[8] C. Pedrinaci, J. Domingue, A. K. Alves de Medeiros, A Core Ontology for Business Process Analysis, in: The Semantic Web: Research and Applications, Springer Berlin Heidelberg, 2008, pp. 49–64.

[9] M. Hepp, D. Roman, An Ontology Framework for Semantic Business Process Management, in: Proceedings of Wirtschaftsinformatik, 2007, pp. 423–440.

[10] C. Di Francescomarino, C. Ghidini, M. Rospocher, L. Serafini, P. Tonella, Semantically-Aided Business Process Modeling, in: The Semantic Web - ISWC 2009, Springer Berlin Heidelberg, 2009, pp. 114–129.

[11] P. Bertoli, F. Corcoglioniti, C. Di Francescomarino, M. Dragoni, C. Ghidini, M. Pistore, Semantic modeling and analysis of complex data-aware processes and their executions, Expert Systems with Applications 198 (2022) 116702. URL: https://www.sciencedirect.com/science/article/pii/S0957417422001804. doi:https://doi.org/10.1016/j.eswa.2022.116702.

[12] S. Eichele, K. Hinkelmann, M. Spahic-Bogdanovic, Ontology-Driven Enhancement of Process Mining With Domain Knowledge, in: Proceedings of AAAI 2023 Spring Symposium on Challenges Requiring the Combination of Machine Learning and Knowledge Engineering (AAAI-MAKE 2023), 2023, p. vvv.

[13] J. Xiong, G. Xiao, T. E. Kalyci, M. Montali, Z. Gu, D. Calvanese, A Virtual Knowledge Graph Based Approach for Object-Centric Event Logs Extraction, in: Revised Selected Papers of the Process Mining Workshops (ICPM-WS 2022), volume 468 of *Lecture Notes in Business Information Processing*, Springer, 2022, pp. 466–478. doi:10.1007/978-3-031-27815-0_34.

[14] A. Swevels, D. Fahland, M. Montali, Implementing Object-Centric Event Data Models in Event Knowledge Graphs, in: Process Mining Workshops, Springer Nature Switzerland, 2024, pp. 431–443.

[15] C. di Ciccio, F. Ekaputra, A. Cecconi, A. Ekelhart, E. Kiesling, Finding Non-Compliances with Declarative Process Constraints Through Semantic Technologies, in: International Conference on Advanced Information Systems Engineering Forum, CAiSE 2019 Forum, Rome, Italy, Springer, 2019, pp. 60 – 74. doi:10.1007/978-3-030-21297-1_6.

[16] S. Khayatbashi, O. Hartig, A. Jalali, Transforming Event Knowledge Graph to Object-Centric Event Logs: A Comparative Study for Multidimensional Process Analysis, in: Proceedings of 42nd International Conference on Conceptual Modeling (ER 2023), Springer-Verlag, Berlin, Heidelberg, 2023, p. 220–238. doi:10.1007/978-3-031-47262-6_12.