

BERMUDA: Towards Maintainable Traceability of Events for Trustworthy Analysis of Non-process-aware Information Systems

Vlad P. Cosma^{1,2}[0000–0001–8022–6402], Thomas T.
Hildebrandt¹[0000–0002–7435–5563], and Tijs Slaats¹[0000–0001–6244–6970]

¹ Copenhagen University, Copenhagen 2200, Denmark
{vco,hilde,slaats}@di.ku.dk

² KMD ApS, Ballerup 2750, Denmark
vco@kmd.dk

Abstract. We propose a method supported by a prototype tool for tracing the relation between domain events and data recorded in non-process-aware information systems via the system and user interfaces. The method contributes to the involvement of domain experts in the tasks of understanding the available data, the domain and the trustworthy and responsible extraction of events from information systems carried out in the first phase of process mining projects.

Keywords: Domain Event · Database · Interface · Information System

1 Motivation

BERMUDA (Business Event Relation Map via User-interface to Data for Analysis) is a method to document and maintain the link between domain events and the data in the information system via the user interface. It facilitates co-operative work spanning the expertise domains of the domain experts using the information system, engineers and data scientists, while still maintaining the necessary barriers in place to protect values, rights and intellectual property. Unlike event abstraction [4] and related work on automatic or semi-automatic event extraction [1,3] we approach the problem in a human-centric manner and make explicit the association between high-level and low-level events, we also do not assume the existence of an event log prior to applying BERMUDA. A similar procedure in [2] focuses on process instance and event level decisions taken when building the event log.

2 Approach

Our method relies on so-called BERMUDA triples **(e,i,d)** recording the relation between respectively a domain event **e**, an interface element **i** of the system in which the domain event is registered in the information system and the location

of the resulting data element **i** in the information system. There are three roles involved in the recording of BERMUDA triples: Data scientist (or analyst), domain expert and system engineer. As guidance towards applying our method we recommend following these steps:

1. **Domain to user interface.** For each domain event **e**, the domain experts record an association **(e,i)** between the domain event **e** and an (user or system) interface element **i**.
2. **User interface to data.** Through code inspection or simulation, system engineers develop the correct database query **d** to extract the data recording the event **e** created via the interface element **i** resulting in a triple **(e,i,d)**.
3. **Triples to event log.** The data scientist merges and refines the database queries and creates the initial version of the event log. The event log entries are enriched with extra attributes that hold a reference to the domain event, the user interface/description and the data source from where the entry originated.

3 Evaluation

In order to qualitatively evaluate the usefulness of the method we conducted two semi-structured interviews, one with a case worker acting as a domain expert and another with a data scientist. From our own experience and the interviews we observed that the method encourages a more structured approach towards event log construction. It also promotes collaboration in order to more efficiently record knowledge gathered from domain experts and system engineers. Both interview participants indicated that the method and tool could be valuable both in training new data scientists and new case workers.

References

1. Baier, T., Rogge-Solti, A., Weske, M., Mendling, J.: Matching of events and activities - an approach based on constraint satisfaction. In: Frank, U., Loucopoulos, P., Pastor, Ó., Petrounias, I. (eds.) *The Practice of Enterprise Modeling*. pp. 58–72. Springer Berlin Heidelberg, Berlin, Heidelberg (2014)
2. Jans, M., Soffer, P.: From relational database to event log: Decisions with quality impact. In: *Business Process Management Workshops. Lecture Notes in Business Information Processing*, vol. 308, pp. 588–599. Springer (2017)
3. de Murillas, E.G.L., Reijers, H.A., van der Aalst, W.M.P.: Case notion discovery and recommendation: automated event log building on databases. *Knowl. Inf. Syst.* **62**(7), 2539–2575 (2020)
4. van Zelst, S., Mannhardt, F., de Leoni, M.: Event abstraction in process mining: literature review and taxonomy, vol. 6, pp. 719–736. Springer New York, New York, NY (2021). <https://doi.org/10.1007/s41066-020-00226-2>