Evaluation of Algorithmic Contributions to Business Process Management

Peer-reviewed author version

DEPAIRE, Benoit; Leopold, H; Mendling, J & Meyerhenke, H (2024) Evaluation of Algorithmic Contributions to Business Process Management. In: Marrella, A.; Resinas, M.; Jans, M.; Rosemann, M. (Ed.). Business process management, BPM 2024, SPRINGER-VERLAG BERLIN, p. 531 -537.

Handle: http://hdl.handle.net/1942/46265

Evaluation of Algorithmic Contributions to Business Process Management

Benoit Depaire[1][0000−0003−4735−0609], Henrik Leopold[2][0000−0003−4862−1829], Jan Mendling[⋆3,4,5][0000−0002−7260−524X], and Henning Meyerhenke[3][0000−0002−7769−726X]

[1] Hasselt University, Hasselt, Belgium
[2] Kühne Logistics University, Hamburg, Germany
[3] Humboldt-Universität zu Berlin, Department of Computer Science, Germany
[4] Vienna University of Economics and Business, Department of Information Systems and Operations Management, Austria
[5] Weizenbaum Institute, Berlin, Germany

Abstract. Algorithms play an important role for business process management. Indeed, all of the 27 papers accepted for BPM 2023 refer to algorithms in one or the other way. Some of these research papers propose new algorithms, such as the paper on the Agent Miner [8], they compose analytical pipelines, such as the paper on process mining video data [5], or they discuss the evaluation of algorithms, such as the paper on sustainability dimensions [3]. So far, however, the BPM community has not developed explicit guidelines for evaluating algorithmic contributions.
In this tutorial, we will describe an overarching methodological framework of how algorithms can be researched. In turn, we discuss its foundations in three dimensions. First, we describe four ontological entities including real-world problem, algorithmic task, algorithm design, and algorithm implementations. Second, we distinguish four types of knowledge, that is a) knowledge relating to tasks and to designs and b) knowledge of and knowledge about. Third, we discuss how methodologically new or better knowledge is established. Fourth, we highlight implications regarding the validity of corresponding research contributions. The objective of this tutorial is to present specific strategies how algorithms can be systematically researched for early career researchers. These are illustrated by good examples of BPM papers.

Keywords: Algorithm engineering · Evaluation of algorithms · Design and Analysis of Algorithms · Process Mining.

⋆ The research of the author was supported by the Einstein Foundation Berlin under grant EPP-2019-524, by the German Federal Ministry of Education and Research under grant 16DII133, and by Deutsche Forschungsgemeinschaft under grant ME 3711/2-1.

1 Outline of the Content

The tutorial builds on our recent framework for researching algorithms [7] with a focus on BPM. We discuss its foundations in ontology, epistemology, and methodology. Different types of contributions faces validity concerns. We highlight specific validity concerns of research on different types of algorithms and corresponding strategies to address them. By clarifying the ideas of the framework based on examples from BPM research, we help early career researchers to express what their research contributes, which in turn helps them to improve the structure of their research and the presentation of their papers.

The first 22 minutes, we will describe the four ontological entities of the framework. These include real-world problem, algorithmic task, algorithm design, and algorithm implementations, which we illustrate using BPM papers. For instance, research by Ali et al (BISE 2024) on waiting times exemplifies the analysis of a real-world problem with algorithmic implications [2]. The specific task of accurately identifying batch activation rules is investigated by Martin et al (TSC 2019) [6]. Algorithm design is the focus of the article on the alpha algorithm by Van der Aalst (TKDE 2004) [1]. Algorithm implementation is a key concern of Dijkman (DPD 2020) [4].

The second 22 minutes focus on epistemology. Four types of knowledge are distinguished: a) knowledge relating to tasks and to designs, and b) *knowledge of* and *knowledge about*. *Knowledge of tasks* relates to *what is* a specific task. *Knowledge about* relates to *what are properties* we know about these tasks. Knowledge of design relates to *what is* a specific algorithm design, while *knowledge about* relates to *what are properties* we know about this algorithm design. Contributions either provide new or better knowledge for each of these four categories. We will discuss examples in each category.

The third 22 minutes discuss methodology, highlighting how new knowledge can be established and justified [7]. New algorithmic tasks and algorithm designs extend the corresponding knowledge categories. New theorems and corresponding proofs establish formal knowledge about algorithms. New empirical results inform empirical knowledge about algorithms. Furthermore, each knowledge type is associated with corresponding research methods: methods for generating knowledge of algorithm designs, methods for generating formal knowledge about Tasks and Designs, and methods for generating empirical knowledge about designs. We will discuss with the participants how each of these contribution types can be carved out in a systematic way.

The fourth 22 minutes investigate nine validity concerns for algorithmic contributions identified by the research framework [7]. These are (1) Ecological validity, (2) Design validity, (3) Implementation validity, (4) Logical validity, (5) Formal correctness, (6) Internal validity, (7) External validity, (8) Construct validity, and (9) Conclusion validity. We will discuss the relevance of each of these validity aspects for different types of contributions, depending on the epistemological and methodological focus.

2 Intended Audience

The intended audience of this tutorial are *doctoral students and early career researchers. No specific prior background knowledge* is required. From our experiences of supervising PhD students, we observe that typical computer science curricula provide limited understanding of epistemological and methodological foundations. With this tutorial, we aim to shed light on the knowledge that is contributed by algorithmic BPM research. To this end, we present concepts from the philosophy of science early-career researchers and specific strategies how algorithms can be scientifically evaluated in a systematic manner.

3 Audience Engagement

We prepare for an audience of up to 60 persons. For that reason, we aim for methods of engagement that scale. At the beginning of the tutorial, we will ask participants to write down for themselves one BPM research papers that spontaneously come to their mind. At each of the four blocks of the tutorial, we will ask participants to fill in a short online questionnaire in which they classify their selected paper according to the different dimensions of the framework. We will show the classification reasons immediately. This will give us the chance to comment on how frequently certain research approaches are considered in BPM research.

4 Short Bio of Presenters

Benoît Depaire is Professor of Business Informatics at Hasselt University, Belgium, and affiliated with the Expertise Centre for Digital Media. His research interests focus on machine learning, statistics, and computational methods to gain business insights from data, as well as the methodology of experimental algorithm design. His research is published, among others, in Business and Information Systems Engineering, Decision Support Systems, Knowledge-Based Systems, Journal of Heuristics, and IEEE Transactions on Services Computing. He is one of the founding organizers of the International Workshop on Event Data and Behavioral Analytics and leads the applied research unit BIARU. See also http://scholar.google.com/citations?user=REbZ-KAAAAAJ

Henrik Leopold is a Full Professor for Data Science at the Kühne Logistics University and a senior lecturer at the Hasso Plattner Institute at the Digital Engineering Faculty, University of Potsdam. Before that, he held positions at the Vrije Universiteit Amsterdam, WU Vienna, and Humboldt-Universität zu Berlin. His research is concerned with the interplay between information systems and business processes, with a focus on leveraging machine learning and natural language processing for process mining, process analysis, and process automation. His research have been published, among others, in IEEE

Transactions on Knowledge and Data Engineering, IEEE Transactions on Software Engineering, Decision Support Systems, and Information Systems. See also http://scholar.google.com/citations?user=qKCTXs4AAAAJ.

Jan Mendling is the Einstein-Professor of Process Science with the Department of Computer Science at Humboldt-Universität zu Berlin, Germany. Before that, he held positions at WU Vienna, Humboldt-Universität zu Berlin, and Queensland University of Technology. His research interests include various topics in the area of business process management, information systems, and process mining. His research has been published, among others, in Management Information Systems Quarterly, IEEE Transactions on Software Engineering, IEEE Transactions on Visualization and Computer Graphics, and Decision Support Systems. He is founding editor-in-chief of Process Science and co-author of the textbooks Fundamentals of Business Process Management and Wirtschaftsinformatik. See also http://scholar.google.com/citations?user= e3LVAMEAAAAJ.

Henning Meyerhenke is Full Professor of Computer Science at Humboldt-Universität zu Berlin since 2018. Prior to that, he held positions at University of Cologne, Karlsruhe Institute of Technology, Georgia Institute of Technology (Atlanta, USA), NEC Laboratories Europe, and University of Paderborn. His research interests revolve around algorithm engineering for large and complex networked systems, in particular regarding the analysis of large complex networks, combinatorial scientific computing, and applied large-scale optimization for the natural sciences. Together with his co-authors, he received the Best Algorithms Paper Award at the 22nd IEEE International Parallel and Distributed Processing Symposium (2008) and the Best Paper Award of the 2015 International Symposium on Foundations and Applications of Big Data Analytics. See also http://scholar.google.com/citations?user=cjr90bEAAAAJ.

References

1. Van der Aalst, W., Weijters, T., Maruster, L.: Workflow mining: Discovering process models from event logs. IEEE transactions on knowledge and data engineering **16**(9), 1128–1142 (2004)
2. Ali, M.A., Milani, F., Dumas, M.: Data-driven identification and analysis of waiting times in business processes: A systematic literature review. Business & Information Systems Engineering pp. 1–18 (2024)
3. Delgado, A., García, F., Moraga, M.Á., Calegari, D., Gordillo, A., Peña, L.: Adding the sustainability dimension in process mining discovery algorithms evaluation. In: International Conference on Business Process Management. pp. 163–177. Springer (2023)
4. Dijkman, R., Gao, J., Syamsiyah, A., Van Dongen, B., Grefen, P., ter Hofstede, A.: Enabling efficient process mining on large data sets: realizing an in-database process mining operator. Distributed and parallel databases **38**, 227–253 (2020)
5. Lepsien, A., Koschmider, A., Kratsch, W.: Analytics pipeline for process mining on video data. In: International Conference on Business Process Management. pp. 196–213. Springer (2023)
6. Martin, N., Solti, A., Mendling, J., Depaire, B., Caris, A.: Mining batch activation rules from event logs. IEEE Transactions on Services Computing **14**(6), 1908–1919 (2019)

7. Mendling, J., Leopold, H., Meyerhenke, H., Depaire, B.: Methodology of algorithm engineering. arXiv preprint arXiv:2310.18979 (2023)
8. Tour, A., Polyvyanyy, A., Kalenkova, A., Senderovich, A.: Agent miner: an algorithm for discovering agent systems from event data. In: International Conference on Business Process Management. pp. 284–302. Springer (2023)