

Organizational Complexity: Insights from Digital Trace Data Research (Extended Abstract)

Bastian Wurm¹

¹LMU Munich School of Management, Ludwigstr. 28, 80539 Munich, Germany

Abstract

This thesis investigates organizational complexity, i.e. the dynamic behavior that emerges from the interaction of distinct parts of an organization. In particular, it examines structural complexity, i.e. how organizations organize themselves as well as *process complexity*, i.e. how they carry out work. In this extended abstract, we focus on the dissertation's contributions to business process complexity that has been conceptualized as the different ways to enact an organizational process. Bridging different streams of Business Process Management (BPM) research, we develop and apply process mining techniques to investigate how business processes complexity changes over time. This thesis offers several contributions to BPM and related fields. First, we discuss how process mining can be used to research organizational processes with digital trace data. Second, we apply process mining techniques to examine how the complexity of organizational processes develops over time. Third, we develop a graph-based measure to operationalize complexity captured in log data. We show that log complexity significantly influences the quality of process models derived with state-of-the-art process discovery techniques.

Keywords

Organizational Complexity, Digital Trace Data, Process Complexity, Business Process Change, Graph Entropy, Process Discovery

1. Motivation

Organizational complexity is the result of the interplay of interdependent parts of an organization [1]. Research in this area has addressed, for instance, the effects of complexity on innovation or organizational performance. Besides, there is an increasing interest in the complexity of organizational processes and how it develops over time (e.g. [2, 3]).

This thesis is motivated by two key observations. First, process complexity is difficult to observe as it involves many parts of an organization. Especially longitudinal studies are scarce, as the problem of observability is multiplied. Second, information systems employed in organizations produce increasing amounts of digital trace data that provide insights into actions performed by organizational actors. It has been argued that this type of data can be used for and might impact our research practices dramatically (e.g. [4, 5]).

To this end, this thesis capitalizes on the opportunities offered by digital trace data to answer the research question: *Which insights do digital trace data provide into organizational complexity and its change over time?*


BPM 2022 Best Dissertation Award, Doctoral Consortium, and Demonstration Resources Track

✉ bastian.wurm@lmu.de (B. Wurm)

🌐 <https://www.en.dmm.bwl.uni-muenchen.de/persons/professoren/wurm/index.html> (B. Wurm)

🆔 0000-0002-1002-5397 (B. Wurm)

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

While the thesis offers contributions to process complexity and structural complexity in organizations, this extended abstract will focus on the studies on process complexity and their contributions to BPM and related fields.

2. Business Process Complexity

We adopt the conceptualization of process complexity as a function of enactments [6, 7]. Following this understanding, processes are networks of interrelated activities. The more activities in the process and the more connections among these exist, the more complex is a given process [7]. To measure process complexity, Haerem and Pentland [7] propose the following formula based on the work of Öser [8]:

$$\sum_g \sum_p ties_{g,p}$$

where p is the number of paths that can be taken to achieve a given goal (g). “The primary understanding reflected in this measure is that task complexity is indexed by the number of paths in the network of events that lead to the attainment of task outcomes” [7, p. 452]. This measure reflects that the complexity of the task system is greater than its single components.

Based on this measure, Goh et al. [3] examine the complexity of scrum software development processes. They find that increasing process complexity reflects requirements for the software. Furthermore, Pentland et al. [2] demonstrate the effect of different variables in digitized processes on process complexity. Surprisingly, they find that digitized processes over time transition from paths with only small variations through *bursts* of complexity before they reach a state in which they exhibit shorts paths with on-going, but limited variation.

It remains unclear how robust and generalizable the simulation findings by Pentland et al. [2] are and whether the key assumptions made hold true in reality. As the authors themselves state in [9], simulation is primarily a tool for hypothesis development warranting further empirical examination and testing. In this thesis, we address the need for further empirical research on process complexity. We detail our contributions in the following.

3. Contributions

3.1. Methodological Contributions

This work entails methodological contributions to process research [10, 11]. We have echoed recent claims regarding the use of digital trace data for research purposes [4, 12]; we proposed and employed process mining [13] as a method to study organizational processes that are supported by information technology.

On the one hand, process mining enables scholars to derive action patterns based on digital action traces and thereby reconstruct and understand how organizational work is carried out. On the other hand, process mining can be employed to test hypotheses on how organizational processes behave and change over time. More specifically, there are three key scenarios of how process mining can be employed to derive insights from digital trace data. (1) Process discovery can be used to derive a process model from digital action traces. (2) Conformance checking can

be applied to compare a process model with enacted process behavior as documented in the log file. (3) Process drift algorithms can be employed to detect changes in how organizational processes are carried out.

Process mining adds to existing procedures and data analysis techniques for process research [14]. Traditional research methods are limited because researchers need to purposefully generate and collect data about organizational phenomena. For example in the case of ethnographic observations, a single scholar can be at one place only and cannot carry out observations indefinitely. Hence, collected data are bound to one place and a limited period. In contrast, process mining allows scholars to investigate organizational processes “in the wild” [15, p. 415] and over extended periods of time. While process mining is useful to detect or test patterns of action, we highlight that process mining should be complemented with interviews to further derive contextual insights and substantiate explanations for how organizational processes behave.

We have applied process mining to investigate how process complexity changes over time. We provided detailed accounts of how we proceeded in our data analysis, specifically how we employed techniques from the process mining of python library [16] to detect and select process variants as well as how our own code complements these techniques. We did this not only for reasons of transparency and replicability, but we hope that future research can draw on these accounts as a means of guidance of how to apply process mining in research projects.

3.2. Contributions to Process Complexity

This dissertation provides contributions to research on process complexity [17]. By investigating process complexity “in the wild” [15, p. 415], we further add to the understanding of how process complexity changes over time. In comparison to Pentland et al. [2], we do not find any indication for *bursts*, i.e. increases of several magnitudes, of process complexity. One possible explanation for this is that activities in organizational processes cannot be arbitrarily combined. Certain activities in organizational processes exhibit logical interdependencies, such that one is required for the other to happen. For example, an invoice cannot be paid before it is created. Furthermore, activities in organizational processes are interdependent, since one activity provides context and signals that influence the enactment of a successive activity [18]. Thus, rather than arbitrarily, activities can be combined according to a Lego analogy: There is only a limited set of combinations of Lego stones (activities) to reach a certain desired figure (an outcome of an organizational process). Hence, in specific situations, a Lego stone cannot be randomly chosen, but one must be chosen that fits with the set of stones already placed.

Additionally, to the process complexity measure proposed by Haerem et al. [7], we suggest two additional measures to operationalize process complexity. We argue that the measure by Haerem et al. [7] quantifies complexity in organizational processes in *absolute* terms. Since this measure is based on the number of total ties, only minimal process variation will lead to a high level of process complexity when a process is often enacted. If variation is stable, a process that is enacted very often will exhibit a higher level of complexity, compared to the very same process enacted once, only. For this reason, we introduce the measure of *relative* complexity that normalizes complexity with respect to the number of enactments. As our applications demonstrate, this distinction is useful, because it allows to differentiate between

the complexity that all process enactments produce and the complexity that each individual enactment contributes. Particularly, in processes with large fluctuations in the number of cases, relative complexity will provide a more nuanced picture.

More generally, this work contributes to building a bridge between Business Process Management and Routine Dynamics. Both fields of study have been referred to as “islands” of process research [11, 19], investigating similar phenomena from different perspectives and with little integration among both fields. By connecting Business Process Management and Routine Dynamics, we contribute to progressing Business Process Management as a behavioral science [11, 20].

3.3. Contributions to Process Mining

We have developed a new measure for process complexity based on graph entropy [21]. While the measure by Haerem et al. [7] has been applied successfully in several studies [2, 17], there are some conceptual issues. An important feature of complex systems is that they dynamically co-evolve and may undergo exponential change. A measurement for complexity needs to adequately represent the complexity of a given entity. However, exponentiality must remain a feature of the entity to be studied, not of the measure itself. As a consequence, we propose a graph-based measure that captures process complexity as it is enacted.

We further apply this measure to investigate the relationship between the complexity of event sequences recorded in event logs and the quality of process models derived with process discovery algorithms. We find that increases in log complexity decrease the quality of discovered process models. More generally, our results point to the importance of the connection between input data and outcomes of process mining algorithms.

Acknowledgments

I would like to acknowledge the continuous guidance and support of my supervisor Prof. Dr. Jan Mendling, without whom this dissertation would not have been possible.

References

- [1] M. Gell-Mann, Complex adaptive systems, in: G. Cowan, D. Pines, D. E. Meltzer (Eds.), *Complexity: Metaphors, models, and reality*, Addison-Wesley, 1994, pp. 17–29.
- [2] B. T. Pentland, P. Liu, W. Kremser, T. Hærem, The Dynamics of Drift in Digitized Processes, *MIS Quartely* 44 (2020) 19–47.
- [3] K. T. Goh, B. T. Pentland, From Actions to Paths to Patterning: Toward a Dynamic Theory of Patterning in Routines, *Academy of Management Journal* 62 (2019) 1901–1992.
- [4] N. Berente, S. Seidel, H. Safadi, Data-Driven Computationally-Intensive Theory Development, *Information Systems Research* 30 (2019) iii–viii.
- [5] D. Lazer, A. Pentland, L. Adamic, S. Aral, A. L. Barabasi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, M. van Alstyne, Life in the network: the coming age of computational social science, *Science* 323 (2009) 721–721.

- [6] B. T. Pentland, M. S. Feldman, Narrative networks: Patterns of technology and organization, *Organization Science* 18 (2007) 781–795.
- [7] T. Hærem, B. T. Pentland, K. D. Miller, Task Complexity: Extending a core Concept, *Academy of Management Review* 40 (2015) 446–460.
- [8] O. A. Oeser, G. O’Brien, A mathematical model for structural role theory: iii, *Human Relations* 20 (1967) 83–97.
- [9] B. T. Pentland, P. Liu, W. Kremser, T. Hærem, Can Small Variations Accumulate into Big Changes?, in: P. Lounsbury, M., Anderson, D.A. and Spee (Ed.), *On Practice and Institution: New Empirical Directions (Research in the Sociology of Organizations, Vol. 71)*, volume 71, Emerald Publishing Limited, 2021, pp. 29–44.
- [10] T. Grisold, B. Wurm, J. Mendling, J. vom Brocke, Using Process Mining to Support Theorizing About Change in Organizations, in: *53rd Hawaii International Conference on System Sciences (HICSS 2020)*, 2020.
- [11] B. Wurm, T. Grisold, J. Mendling, J. vom Brocke, Business Process Management and Routine Dynamics, in: *Cambridge Handbook of Routine Dynamics*, Cambridge University Press, 2021, pp. 513–524.
- [12] A. Lindberg, Developing Theory through integrating Human & Machine Pattern Recognition, *Journal of the Association for Information Systems* (2019).
- [13] W. M. van der Aalst, *Process mining: Data Science in Action*, 2016.
- [14] A. Langley, Strategies for Theorizing from Process Data, *The Academy of Management Review* 24 (1999) 691–710.
- [15] A. Parmigiani, J. Howard-Grenville, Routines revisited: Exploring the capabilities and practice perspectives, *Academy of Management annals* 5 (2011) 413–453.
- [16] A. Berti, S. J. van Zelst, W. M. van der Aalst, Process mining for python (PM4py): Bridging the gap between process- and data science, in: *Proceedings of the ICPM Demo Track 2019, co-located with 1st International Conference on Process Mining (ICPM 2019) 2019*, 2019, pp. 13–16. [arXiv:1905.06169](https://arxiv.org/abs/1905.06169).
- [17] B. Wurm, T. Grisold, J. Mendling, J. vom Brocke, Measuring Changes of Complexity in Organizational Routines, in: *Proceedings of the Eighty-first Annual Meeting of the Academy of Management*, 2021.
- [18] W. Kremser, B. T. Pentland, S. Brunswicker, Interdependence within and between routines: A performative perspective, in: M. Feldman, L. D’Adderio, P. Jarzabkowski, K. Dittrich (Eds.), *Research in the Sociology of Organizations. Routine dynamics in action: replication and transformation*, volume 61, 2019, pp. 79–98.
- [19] B. T. Pentland, E. Vaast, R. Wolf, Theorizing Process Dynamics with directed Graphs: A Diachronic Analysis of digital Trace Data, *MIS Quartely* (2021) 967–984.
- [20] J. Recker, J. Mendling, The State of the Art of Business Process Management Research as Published in the BPM Conference: Recommendations for Progressing the Field, *Business & Information Systems Engineering* 58 (2016) 55–72.
- [21] A. Augusto, J. Mendling, M. Vidgof, B. Wurm, The connection between process complexity of event sequences and models discovered by process mining, *Information Sciences* (2022).