

Towards Routines Mining – Designing and Implementing the Argos Miner, a Design Science Artifact for Studying Routine Dynamics with Process Mining

Pascal Budner
University of Cologne
budner@wiso.uni-koeln.de

Bastian Wurm
Vienna University of
Economics and Business
bastian.wurm@wu.ac.at

Christoph Rosenkranz
University of Cologne
rosenkranz@wiso.uni-koeln.de

Jan Mendling
Humboldt University of
Berlin
jan.mendling@hu-berlin.de

Abstract

Digital artifacts increasingly support actors in carrying out organizational routines. These artifacts leave digital trace data, that is, time-stamped data about what actions actors performed. While extant research on routines largely builds on qualitative methods, the increasing ubiquitousness and prevalence of trace data enable novel methodological opportunities. However, several challenges currently hinder the adoption of trace data in empirical research on routines in general and their dynamics in particular. Promising approaches such as process mining are neither designed for nor sensitive to the concept of routines. In this paper, we follow a design science research approach to develop the first iteration of an artifact, which we coin Argos Miner. This artifact is based on process mining algorithms and overcomes challenges inherent in adopting process mining in routine dynamics research. It enables scholars to capture reality in flight by analyzing routine dynamics using a computational, mixed-methods approach.

1. Introduction

Organizational routines (hereafter, just routines) are “repetitive, recognizable patterns of interdependent actions, carried out by multiple actors” [1]. They are at the core of accomplishing work in organizations [1]. Digital artifacts (i.e., man-made objects with computing abilities) play an important role for routines, as they constrain and enable their performance and their understanding [2, 3]. Routines are oftentimes supported by or completely carried out by information technology (IT). When IT is used, digital trace data are generated. While the study of routines mainly relies on qualitative methods, such as ethnography or interviews, the increasing availability of digital trace data provides tremendous opportunities for theory development and theory testing about how organizational work is carried

out [4, 5]. In this light, scholars have called for building on this kind of data and proposed to computationally study routines by use of process mining algorithms developed in the business process management (BPM) community [6, 7].

While there are first studies that have adopted process mining to investigate and theorize about routines [8, 9], several challenges currently hinder the broader uptake of process mining as means to study routines. In particular, process mining algorithms are not sensitive to the study of the *dynamics* of routines [7]. Dynamics of routines only become visible in the light of time and at very high levels of granularity. However, process mining algorithms are mainly static and often treat infrequent behavior as outliers [10]. Furthermore, the application of process mining algorithms requires considerable technical expertise posing a limitation to the adoption of process mining for routine dynamics scholars [9]. To address these limitations, we ask the following research question (RQ):

RQ: *How can we adopt algorithms from process mining to computationally study routine dynamics?*

We apply a design science research approach [11] to answer this research question. To this end, we propose and implement a software artifact that enables scholars to readily investigate routine dynamics with digital trace data and process mining. Thereby, our contribution is two-fold.

First, we propose design principles for a class of what we call *routines mining artifacts*. These artifacts sensitize and adapt selected process mining algorithms for the study of routines, and combine them in one system. To this respect, our study serves as a bridge connecting the two separated “islands” of process research, namely routine dynamics and business process management [9, 12].

Second, we provide a concrete instantiation of such an artifact, the so-called *Argos Miner*. This artifact

allows scholars to readily investigate routines from different perspectives employing digital trace data without requiring much technical knowledge of process mining. In contrast to existing commercial process mining tools, our artifact leverages a selected set of algorithms that, as we argue, align well with the assumptions of routine dynamics research.

2. Theoretical Background

2.1. Routine Dynamics

Organizational routines are central to performing work in organizations and have been defined as “repetitive, recognizable patterns of interdependent actions, carried out by multiple actors” [1]. Examples of routines are the Sushi preparation routine [13], the patient treatment routine [9], or the purchase-to-pay routine [8]. Routines have long been regarded as stable entities [14]. With the practice-based turn, the view on routines has shifted. Grounded in strong process theory [16], routine dynamics emphasizes the dynamic and emergent properties of routines [15]. Fundamental to this contemporary perspective on routines is the distinction and interaction between the ostensive and performative aspects [1].

The *ostensive* aspect refers to the mental model that actors hold about a specific routine; it thus guides actors' routine enactment. The *performative* aspect, in comparison, captures “specific actions at specific points in time carried out by specific actors” [1]. Through the interaction of both aspects, routines are continuously changing, but remain recognizable, a characteristic that is also described in the paradox of the (n)ever-changing world [16]. The field of routine dynamics is mainly interested in how and why routines change over time. For example, Dittrich et al. [17] show that actors reflect on how they perceive and enact the routines, leading to change thereof.

Artifacts play an important role for almost all routines, whether it is the Sushi preparation routine aided by Sushi knives [13] or the Purchase-to-Pay routine supported by enterprise resource planning (ERP) systems [8]. Routines and respective artifacts demonstrate a reciprocal relationship. The artifact enables and constrains the performative aspect and the ostensive aspect of routines and vice versa [2]. Increasingly, these artifacts are digital, that is, they are endowed with information processing capabilities. For example, a study by Berente et al. [18] explains how NASA introduced an integrated information system to manage their various research centers. They find that

routines can serve as “shock absorbers” and help mitigate the misalignments between the artifact (the information system), as well as the ostensive and performative aspects of routines.

2.2. Computational-driven Theorizing

When organizational work is supported by digital artifacts, often some form of digital trace data is recorded. Digital trace data comprises information about what actions actors perform at a given point in time [4]. The increasing penetration of IT in work and everyday life also increases the availability of this type of data. In this context, a growing number of scholars [4, 5, 19] argue that this type of data can spur methodological innovations contributing to the testing of existing theories or the development of new theory, altogether.

Recent articles [e.g., 4, 5, 19] propose the combination of human and machine reasoning, in the spirit of mixed-methods approaches, to test or develop theory. On the one hand, computational artifacts can be used to inductively identify or deductively test patterns in digital trace data. On the other hand, scholars need to make sense of the data by mapping found patterns to an existing discourse or lexicon [4] or by discovering patterns themselves. Research highlights that contextual information can be used to validate computationally derived patterns, for example, through interviews or observations [4, 19]. Figure 1 outlines such an approach based on the work by Lindberg [5]. In our case, the routines mining artifact serves as the machine sensemaking device.

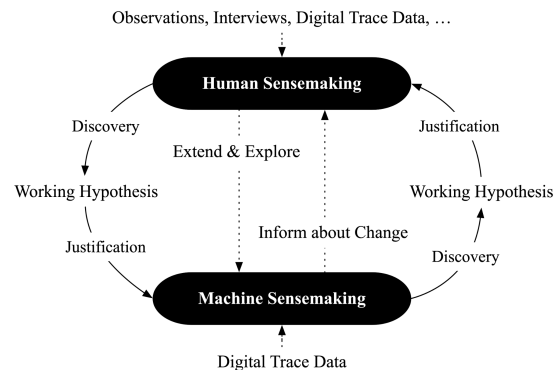


Figure 1. A computational-driven approach to study routine dynamics (based on [5])

2.3. Studying Routines with Process Mining

To date, research on routines largely rests on qualitative methods such as interviews or ethnography

[20]. As discussed in the previous section, with the increasing intertwining of organizational work and IT, an increasing amount of digital trace data is produced that provides enormous potential for computational-driven theory development and testing [4, 5]. In the context of routines, multiple authors have argued that process mining [10] can be utilized to make visible, analyze, and theorize routines using digital trace data [6, 8, 9].

Process mining [21] comprises a set of techniques to analyze business processes and their performance [22] and is mainly researched in the BPM community. Over the last years, process mining has received tremendous uptake in practice [23]. Three types of process mining algorithms are particularly useful for the analysis of routines [6]: process discovery, concept drift, and conformance checking.

First, *process discovery* [24] can be employed to visualize action patterns of routines based on the recorded digital action traces. The resulting process model provides detailed insights into how a routine is performed. Based on process discovery, researchers can zoom in and out [25] on resulting process models to understand the behavior of routines at different levels of granularity [7]. Second, *concept drift algorithms* [26] point to changes in routines over time. For example, [27] propose a sliding window approach to test for any differences in process behavior. This helps routine scholars to see when a routine changes. Third, *conformance checking algorithms* [28] compare an event log with a process model. Researchers can employ this class of algorithms, for example, to hypothesize and test the effect of routine redesign [7].

Some studies have employed process mining to research and have theorized about routines based on digital trace data. For example, Wurm et al. [8] use process mining to analyze routines that are supported by ERP systems. The authors extracted event logs that cover the performance of two core routines: the purchase-to-pay and the order-to-cash routines for two countries each. Based on these data, they find that complexity in all investigated routines fluctuates, but they do not find any indications for “bursts” of complexity as described by Pentland and others in a recent simulation study [7]. To compute their measures and carry out their analysis, the authors relied on the Process Mining for Python library (PM4Py) [29] but had to implement the functionalities to compute the different process complexity largely manually by themselves.

Another example is a study using data from hospital systems to analyze how the patient care routine at a dermalogic hospital changes over time by

Pentland et al. [9]. The authors find that during flu season the routine changes, but in such subtle ways that even personnel at the hospital were not aware. Like the study by Wurm et al. [8], Pentland et al. [9] were required to implement functionalities for large parts of their analysis manually themselves.

Despite these initial applications and the general potential of process mining for the study of routine dynamics, there are currently four key limitations that hinder its broader adoption by the research community.

First, business process management and routine dynamics are largely unconnected “islands” of process research [12], following different “ends-in-views”, pursuing different knowledge claims, and adopting different terminologies [9, 12].

Second, methodological and theoretical challenges arise when working with digital trace data. Digital trace data “are not given but produced” [19]. To overcome these challenges, Osterlund et al. [19] offer various principles that researchers working with digital trace data should follow. For example, trace data should be investigated from multiple perspectives through zooming in and out [25].

Third, many process mining techniques require the availability of high-quality event logs [21] that capture the complete enactment of routines in an end-to-end fashion. However, the enactment of routines often requires the usage of multiple application systems, each supporting specific actions of the routine. The resulting digital trace data is thus not stored in one application system only, but distributed over multiple systems. Combining these data into one integrated event log can pose a major technical challenge [6].

Fourth, process mining algorithms are primarily built for post-hoc analysis, restricting the researcher to sense-making at hindsight [7]. Consequently, researchers are not able to collect contextual information, such as interviews or ethnographic observation [6], while change is ongoing.

To study routine change as it unfolds, we need a methodological tool set that is “sensitive to the phenomena we are trying to understand” [7]. In the following, we address the outlined limitations by developing design principles for a class of artifacts, which we label routines mining artifacts, and instantiate an open-source prototype thereof.

3. Method

We employ the design science research (DSR) framework proposed by Vaishnavi and Kuechler [30] to design, instantiate, and evaluate a design science artifact for studying routine dynamics with process

mining. Our DSR approach entails five iterative activities (see Figure 2).

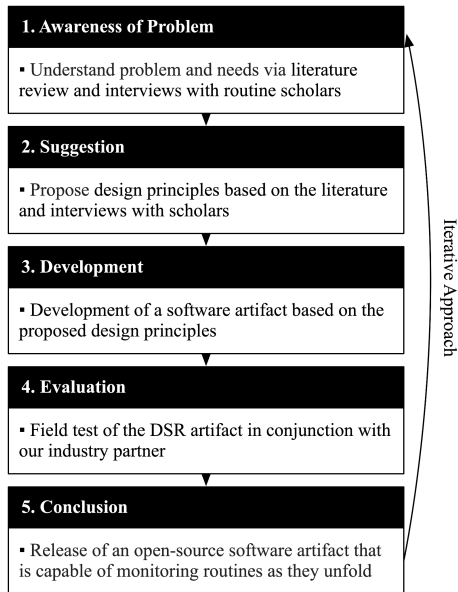


Figure 2. Our iterative DSR approach

First, we started with the problem awareness activity. In this activity, we consulted the extant literature on routine dynamics, process mining, and digital trace data, and conducted two semi-structured interviews with target users, which are routine dynamics scholars. Both helped us to ensure the relevance of our design science research approach with regards to the problem domain [31].

Second, in the suggestion activity, we developed the first iteration of design principles for our software artifact. Design principles are normative design decisions that guide how the artifact should be instantiated to fulfill the requirements [32]. We followed Gregor et al.’s [32] guidelines to specify the design principles.

Third, in the development activity, we instantiated the DSR artifact in the form of a software artifact. In that regard, we used Python as our programming language as it offers a rich ecosystem of open source libraries suitable for our purposes. Building on our proposed design principles, we made use of the open-source process mining library PM4Py [29] and extended it to our needs. PM4Py is considered a state-of-the-art research platform in Python for process mining that offers a rich variety of built-in process mining algorithms. Other open-source libraries, such as bupaR, offer in their current state fewer functionalities and are not part of the Python ecosystem.

Fourth, we evaluated the instantiated artifact “in the wild” [33] to provide a “proof-by-demonstration” [34]. For doing so, we collaborate with a German

insurance company (hereafter, INSUR). INSUR is part of one of the largest multinational insurance organizations in the world and generates yearly revenues of about 10 billion Euro with approximately 9,000 employees distributed over several sites in Germany. Specifically, we utilized the access to INSUR to install the software artifact within INSUR’s IT landscape. Thereby, we were able to instantiate and test the artifact in a real-world setting, with real digital artifacts and digital trace data. Until now, we have collected approx. 20GB of trace data spanning multiple routines, over a time of two months.

Finally, with the paper at hand, we conclude a first iteration of the DSR approach by presenting our results and releasing the first version of our software artifact. The resulting artifact is available as open-source (GNU General Public License v3.0) software on GitHub. While this activity closes the DSR approach, we want to emphasize that the released DSR artifact will continue to evolve over time (e.g., to fix malfunctions or to include new features based on feedback and gained experience). Specifically, we envision a broader evaluation that covers experiences beyond the expertise of the author team.

4. Designing a Routines Mining Artifact

In this section, we present our conceptualization for a routines mining artifact (i.e., an artifact that utilizes process mining to support studying routine dynamics in a computational, mixed-methods approach; see Section 2). This has several implications as it requires an artifact that is (1) aligned with the assumptions and goals of routine dynamics research [12], and (2) supports the longitudinal study of routine dynamics while capturing reality in flight [35].

The following conceptualization is structured by the idea that the artifact itself – as every digital artifact – is in its nature an information processing system (i.e., it processes inputs into outputs). Therefore, we start with the *input layer* (i.e., the data collection). Then, we continue with the *processing layer* (i.e., the analysis of the routines). Finally, we present the *output layer* (i.e., the presentation of the results of the analysis). We follow the guidelines proposed by Gregor et al. [32] for the specification of the design principles (DP).

Input Layer. Process mining requires data in form of event logs (i.e., including timestamps, correlation IDs, and activity names) [10], and process mining algorithms often presuppose certain data structures [e.g., 36]. Additionally, when routine performances are supported by digital artifacts, they often span multiple

artifacts that might have distinct formats for their event logs. To overcome those challenges during data collection, we derive three design principles covering the input layer (see Table 1 for an overview).

Table 1. Input layer design principles

Aim	Mechanism	Rationale
DP-I.1. Principle of Real-time Collection		
To address the demand for appropriate data	<i>Continuously</i> extract data from sources that emit event data about routine performances	Examining routines as they unfold requires data capturing ongoing performances [35]
DP-I.2. Principle of Semantic Heterogeneity		
To load data coming from diverse data sources	Make the data input format configurable	Routines spanning multiple artifacts might emit their events in distinct formats
DP-I.3. Principle of Data Abstraction		
To be able to handle large amounts of continuously emitted event data	Transform and persist event data in an abstracted representation [36]	Long-term storage of abstracted data representations relaxes issues with storage sizing [10]

DP-I.1. Principle of Real-time Collection.

Routine dynamics is concerned about how and why routines stabilize or change over time [15]. This requires analytical tools that help to understand how routines evolve *as* they evolve [35]. Consequently, the routines mining artifact needs appropriate data that capture routine performances as they unfold. Accordingly, we propose the first principle: Continuously extract event data from all sources that emit data about routine performances.

DP-I.2. Principle of Semantic Heterogeneity.

Event data might be spread over multiple digital artifacts that each might have distinct event log formats (e.g., JavaScript Object Notation, JSON, or comma-separated values, CSV). Hence, to load data coming from diverse data sources, the artifact needs to be able to parse different input formats. This yields the second design principle: Provide options to configure the log format for each input source.

DP-I.3. Principle of Data Abstraction. The longitudinal study of routine dynamics continuously generates increasing amounts of data (theoretically, infinite amounts of data depending on the duration of the research study). This bears several challenges, such as for data storage and data loading. The field of stream-based process mining suggests storing event data in abstracted data representations that relaxes

these issues [36]. Hence, to be able to handle large amounts of event data, we propose a third principle for the input layer: Transform and persist event data in an abstract data representation [e.g., 36].

Processing Layer. Process mining per-se is not made to study routine dynamics [12]. Instead, process mining is primarily used to discover, monitor, and improve business processes in practice [37]. This has implications when using process mining for routine dynamics research. For instance, some algorithms assume that event logs contain outliers (or noise) that need to be ignored when discovering processes, such as the heuristic mining algorithm that ignores infrequent behavior [10]. This is in stark contrast to routine dynamics, which assumes varying behavior as indications of the underlying phenomena (i.e., routine dynamics) [38] or as “normal daily variations” [9]. As a result, we propose the following three design principles covering the processing layer to overcome hurdles when utilizing process mining to study routine dynamics (see Table 2 for an overview).

DP-P-1. Principle of Incremental Calculation.

Routines are emergent accomplishments that have a continuous tension between stability and change [1, 15]. Additionally, varying behavior is inherent in routine dynamics [38]. Hence, to make sense of routine dynamics, we need to capture them as they evolve [35] while acknowledging infrequent behavior. This requires online process mining algorithms that incrementally ingest emitted event data without removing infrequent behavior. Hence, we posit the first design principle for the processing layer: To enable the analysis of routines as they unfold, the artifact needs to make use of online process mining algorithms that acknowledge infrequent behavior [e.g., 36].

DP-P-2. Principle of Algorithmic Configurability.

Change depends on context [9]. For some routines, even small changes are notable deviations (e.g., in vaccine production routines), while for other routines small deviations are normal (e.g., in electronic medical records [9]). Hence, the artifact needs to account for different perspectives. This yields the second design principle: To provide multiple perspectives on routine dynamics, make the applied algorithms adjustable to the context’s needs.

DP-P-3. Principle of Alerting. Routine enactment is distributed over space, time, and actors [15]. Scholars, however, can only observe what is in their point of view [39]. The computational, mixed-methods approach introduced in section 2 bears the potential to relax this issue. In particular, the artifact could actively pinpoint scholars through alerts to notable dynamics, for instance, by sending an email as soon as a change is

detected. This enables scholars to sharpen their focus and timely make sense of computationally identified patterns. Consequently, we posit: To notify scholars that a notable routine dynamic is unfolding, the artifact needs to actively inform scholars about computationally identified dynamics.

Table 2. Processing layer design principles

Aim	Mechanism	Rationale
DP-P.1. Principle of Incremental Calculation		
To enable the analysis of routine dynamics as they unfold	Make use of online process mining algorithms that acknowledge infrequent behavior (e.g. [36])	Online algorithms incrementally ingest data and do not require all data at anytime [36]
DP-P.2. Principle of Algorithmic Configurability		
To provide multiple perspectives on routine dynamics	Make the parameters of the applied algorithms adjustable and sensitive to the context's needs	What is perceived as change depends on the context [9]
DP-P.3. Principle of Alerting		
To notify scholars that a notable routine dynamic is unfolding	Actively pinpoint scholars through alerts (e.g., e-mails) to notable changes	Scholars need to make sense of patterns as soon as they unfold [5, 35]

Output Layer. Studying routine dynamics is inherently complex as they are “generative systems, distributed in time and space, and enacted by multiple participants” [39]. As Pentland and Feldman argue, any empirical research on routine dynamics needs to identify and compare routines and their performances, preferably over time [39]. Hence, a routines mining artifact needs to facilitate those activities and support the researcher to conduct the computational, mixed-methods approach as illustrated in section 2. In that regard, we derive three design principles for the output layer (see Table 3).

DP-O.1. Principle of Temporal Traceability. As routines are distributed over time, space, and actors, they have an inherent potential to change over time. Hence, a routines mining artifact needs to shed light on the temporal evolution of the routine in general and the constituting relations of its actions in particular to allow sensemaking of its dynamics. Several authors have suggested that directed graphs are helpful in that regard as they situate actions in relation to each other and [9, 40]. In addition, to enable scholars to trace the unfolding of routines over time, the artifacts require time-variant directed graphs. For instance, the artifact should offer the option to adjust the time window used to visualize the graph. In sum, we propose: Visualize

routines as time-variant directed graphs to facilitate temporal traceability.

DP-O.2. Principle of Selection Flexibility. While the unit of observation in routine dynamics research is situated action, the unit of analysis is patterns of actions [38]. Hence, to fully make sense of the unfolding of the routine, the artifact needs to provide information on different levels of abstraction and with different points of interest, e.g., by zooming in and out [25]. For instance, on the one hand, the artifact could provide information on the action level, such as how frequently a handoff from one action to another action has been enacted within a routine [41]. On the other hand, the artifact could provide information on the routine level, such as its complexity [7], or even on the cluster-level of routines [42]. To conclude, we propose: Provide visualizations and statistics on different levels of abstraction.

DP-O.3. Principle of Reproducibility. In contrast to process mining, which primarily focuses on improving processes in practice [21], the routines mining artifact's purpose is to facilitate computational theorizing [4]. To enable scholars to build theory, the artifact needs to provide explanations and references for the applied algorithms to increase the transparency and reproducibility of the research. Accordingly, we derive the following principle: Provide explanations for the applied algorithms to avoid a black box.

Table 3. Output layer design principles

Aim	Mechanism	Rationale
DP-O.1. Principle of Temporal Traceability		
To make routines and the relations of their actions visible over time	Visualize routines as time-variant directed graphs	Directed graphs situate actions in relation to each other at certain time slices, which facilitates traceability [9, 40]
DP-O.2. Principle of Selection Flexibility		
To enable scholars to make sense of the data	Provide visualizations and metrics on different levels of abstraction	Scholars need to be able to engage with the tool for sense-making (e.g. by zooming in & out [25])
DP-O.3. Principle of Reproducibility		
To enable scholars to develop theory	Provide explanations (e.g., literature references) for the applied algorithms	Providing explanations and references for the applied algorithms increases the transparency

5. Instantiation of the Argos Miner

In this section, we outline a prototypical instantiation of a routines mining artifact following our proposed design principles. We implemented the artifact, which we coin Argos Miner¹, in Python building on the open-source library PM4Py [29]. While PM4Py offers several built-in algorithms for analyzing event logs, current support for streaming analysis, interactive visualizations, and data ingestion is limited. This required us to develop these features ourselves. In line with the previous section, we demonstrate the implementation details along the three layers.

Input Layer. The Argos Miner provides a configuration file to specify data sources and their corresponding event log format (*Principle of Semantic Heterogeneity*). To date, the Argos Miner supports CSV and JSON as data structures and files as well as Apache Kafka as input sources. In our field test, we made use of a file input formatted in CSV.

Next, the Argos Miner continuously extracts event data from the defined input sources (*Principle of Real-time Collection*), parses them, and stores them in a stream-based abstract representation (S-BAR) [36] (*Principle of Data Abstraction*). As soon as new data is added, the Argos Miner automatically processes these data and updates its models.

Processing Layer. The Argos Miner then makes use of an online process discovery algorithm [36] (*Principle of Incremental Calculation*) to derive directly-follows relations between actions over time. If novel directly-follows relations are detected, the Argos Miner sends an e-mail to inform a specified recipient list about the routine change (*Principle of Alerting*). Additionally, the Argos Miner provides functionality to compare a routine and its directly-follows relations via adaptive time windows (*Principle of Algorithmic Configurability*).

Output Layer. The visualization of the Argos Miner can be divided into two parts. The first part of the output layer (upper part of Figure 3) depicts overall information regarding the collected digital trace data. In our case, we mined 21,399,840 events, 15,498 distinct directly-follows relations as well as 690 distinct actions covering two months of routine performances. Additionally, a line chart illustrates the overall distribution of emitted events over time, indicating in our case that INSUR has a lower activity

of routine performances on weekends (see the regular dips in the line chart).

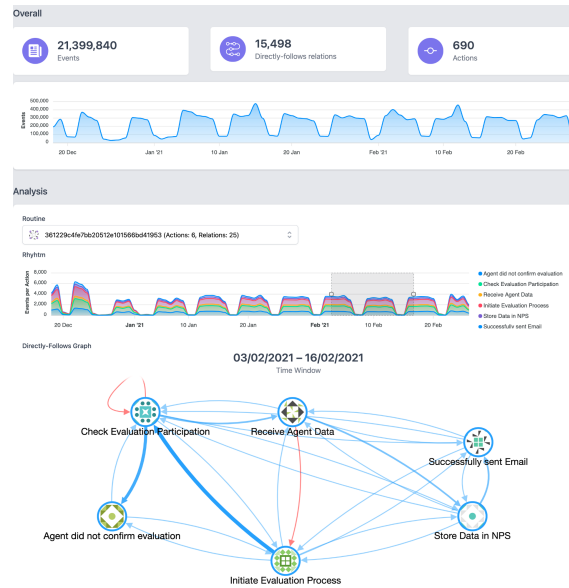


Figure 3. Insurance agent evaluation routine

The second part of the output layer (lower part of Figure 3) contains the option to select a mined routine. After selecting a routine, the corresponding activity frequency graph, as well as the directly-follows graph (DFG), is shown. The activity frequency graph represents the frequencies of each activity over time for the entire observation timespan (*Principle of Temporal Traceability*). To ease the differentiation between activities, the frequency of each activity is depicted with a different color in the line chart. In the DFG the frequency of the directly-follows relation is represented by the thickness of the line. While the default DFG is derived based on the complete timespan, users can use the activity frequency graph to select a timespan and re-instantiate the directly-follows graph (*Principle of Selection Flexibility*). If the selection procedure is repeated, one can contrast the DFGs of different time intervals: While novel actions and directly-follows relations are visualized in green, those that are absent (compared to the previous timespan) are highlighted in red. Thereby, this visualization sheds light on how the routine evolved over time.

Finally, all algorithms employed in the Argos Miner are publically available (including references to the literature) from our GitHub repository. Thus, the employed algorithms and their implementation can be verified (*Principle of Reproducibility*).

These visualizations can now be used to derive insights about the mined routines. For instance, Figure 3 shows the insurance agent evaluation routine of INSUR consisting of six activities. The activity

¹ According to Greek mythology, Argos was a giant having myriad eyes with which he could observe nearly everything. After his death, his eyes were preserved in a peacock's tail.

frequency graph indicates a highly rhythmic routine as the frequencies of all actions in this routine seem to be synchronized. Again the graph indicates low activity on weekends. Further, one can see that while for the selected timespan (03/02/2021 - 16/02/2021) all actions occurred, two directly-follows relations were absent in comparison to the full timespan. In contrast, the credit rating routine (see Figure 4) consists of more activities (15) and seems to be more chaotic as the activity frequency graph is not as synchronized as in Figure 3. Additionally, one can see that for the selected timespan several directly-follows relations, as well as one activity, did not occur. After talking to corresponding human actors participating in the routine, we found out that the routine depicted in Figure 3 is highly automated, whereas the credit rating routine is a rather manual and more complex routine consisting of several conditional execution paths. In a more detailed analysis, one could, for example, further examine whether automated work is generally more rhythmic than manual work. Due to space constraints, we leave this investigation for future research.

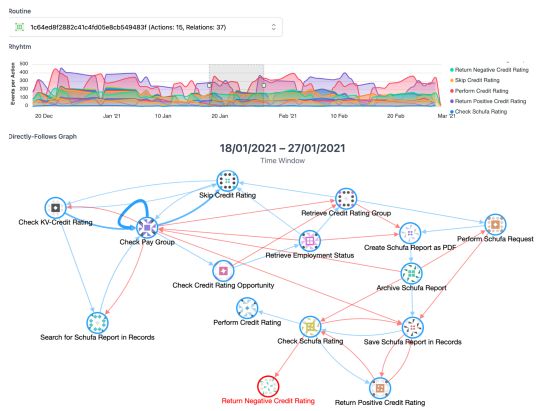


Figure 4. Credit rating routine

6. Discussion

6.1. Implications for Research

In this paper, we have presented a design science research study in which we proposed the class of routines mining artifacts to computationally study routine dynamics. Further, we have developed a concrete instantiation of such an artifact that we coined Argos Miner. The artifact allows scholars to study routines as they unfold, from different angles, and at scale. There are three key contributions of our work that we want to highlight.

First, with this work, we pave the way for a computational-driven theorizing approach (as elaborated in Section 2.2) in routine dynamics

research. This approach capitalizes on the increasing availability of digital trace data [4] to complement traditional research methods, such as ethnography and interviews. In contrast to traditional research practices in routine dynamics research, computational-driven theorizing enables scholars to examine routines across time, space, and at a large scale (e.g., all routines in an organization that emit trace data). Additionally, the rise of computational actors in organizational work (e.g., algorithmic trading agents), will require scholars to employ computational tools to study resulting digitized routines.

Second, with the Argos Miner, we provide a concrete instantiation of a routines mining artifact that routine scholars can readily use. The Argos Miner bundles several features to analyze routine dynamics based on digital trace data in one application. Up to now, to analyze routines, scholars have needed to implement many of the desired functionalities themselves [8, 9]. On the other side of the spectrum, researchers may have used commercial software to analyze and theorize about routines. Yet, these programs are built for practice and not research, emphasizing practical needs over research requirements. Thus, results derived from commercial tools are difficult, if not impossible, to validate. The Argos Miner addresses these limitations by offering tailored functionality for routine dynamics research in an open-source application. Compared to commercial software, all of the implemented algorithms in the Argos Miner align with the assumptions of routine dynamics and are white-boxed, contributing to transparent and replicable research.

Third, our work contributes towards building a bridge between the largely isolated “islands” of process research: BPM and routine dynamics [9, 43]. As previously argued [9, 43], both research fields can profit tremendously from one another. With this paper, we take a step towards the convergence of both fields.

6.2. Implications for Practice

While our work primarily addresses scholars, it also offers implications for practice. With the Argos Miner, practitioners are equipped with a ready-to-use tool to monitor how their organizational routines evolve over time. For instance, the Argos Miner enables practitioners to track whether intended changes are reflected in routine performances. On the other hand, practitioners may decide to receive alerts if their routines deviate from previous performances to timely circumvent unwanted changes in critical routines, such as in highly standardized production routines.

6.3. Limitations and Future Work

There are several limitations to our work at hand. First, in this paper, we report our results from the first iteration of our DSR approach. While we demonstrated a prototypical instantiation, the Argos Miner will be subject to an evaluation beyond the expertise of the author team. We intend to do so in collaboration with our case organization.

Second, our current implementation assumes that all activities are ingested in the correct order, i.e. activities are correctly sorted by time. However, some artifacts may produce trace data slower than others. Where this is the case, mined directly-follows relations may be compromised and the resulting DFG will not accurately represent routine enactment. We plan to relax this assumption by implementing a decoupling queue that temporarily stores digital traces and orders them before they are loaded.

Third, while we have implemented the possibility to monitor change of routines, we plan to implement further features that allow for more granular analysis of routine dynamics. For example, we aim to implement a measure by Augusto et al. [44] to monitor how complexity in routines changes over time.

At the same time, there remain plenty of opportunities regarding the development of algorithms to investigate routines and change thereof. For example, there is an opportunity to develop algorithms that measure variability and change of routines on very abstract as well as on more granular levels. Together, these measures could provide valuable information on whether and how routines change in the light of time. More generally, the discourse in the routine dynamics community points to a plethora of fruitful research endeavors for BPM, and process mining in particular. While BPM research often assumes processes to be static, the processual nature of routine dynamics opens up several research opportunities.

Finally, while digital trace data offers tremendous research opportunities, several issues should be accounted for when employing this type of data. First and foremost, researchers need to have access to an organization's trace data from start to end to fully reflect routine performances. This might bear issues regarding data access, privacy, and completeness that one needs to solve. Second, digital trace data comes with potential validity issues [19]. To account for these issues, researchers should validate the trace data they employ. Finally, we agree with others [6, 43], who argue that digital trace data analyses should be complemented by additional data (see section 2.2.)

7. Conclusion

In this paper, we introduced the class of routines mining artifacts and presented the Argos Miner, a concrete instantiation thereof to computationally study routines using digital trace data. The Argos Miner readily bundles several functionalities from process mining in one system and adjusts them to the context of routine dynamics research. We expect that our artifact will support future routine dynamics research in theory development as well as in theory testing.

8. References

- [1] Feldman, M.S., and B.T. Pentland, "Reconceptualizing Organizational Routines as a Source of Flexibility and Change", *Administrative Science Quarterly* 48(1), 2003, pp. 94.
- [2] D'Adderio, L., "Artifacts at the centre of routines: performing the material turn in routines theory", *Journal of Institutional Economics* 7(2), 2011, pp. 197–230.
- [3] Pentland, B.T., and M.S. Feldman, "Designing routines: On the folly of designing artifacts, while hoping for patterns of action", *Information and Organization* 18(4), 2008, pp. 235–250.
- [4] Berente, N., S. Seidel, and H. Safadi, "Research Commentary—Data-Driven Computationally Intensive Theory Development", *Information Systems Research* 30(1), 2019, pp. 50–64.
- [5] Lindberg, A., "Developing Theory Through Integrating Human and Machine Pattern Recognition", *Journal of the Association for Information Systems*, 2020, pp. 90–116.
- [6] Grisold, T., B. Wurm, J. Mendling, and J. Vom Brocke, "Using Process Mining to Support Theorizing About Change in Organizations", (2020).
- [7] Pentland, B.T., P. Liu, W. Kremser, and T. Haerem, "The Dynamics of Drift in Digitized Processes", *MIS Quarterly* 44(1), 2020, pp. 19–47.
- [8] Wurm, B., T. Grisold, J. Mendling, and J. Vom Brocke, "Measuring Fluctuations of Complexity in Organizational Routines", *Proceedings of the Eighty-first Annual Meeting of the Academy of Management*, In Sonia Taneja (Ed.) (In Press).
- [9] Pentland, B.T., E. Vaast, and J.R. Wolf, "Theorizing Process Dynamics with Directed Graphs: A Diachronic Analysis of Digital Trace Data", *MIS Quarterly* 45(2), 2021, pp. 967–984.
- [10] van der Aalst, W., *Process Mining*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2016.
- [11] Hevner, March, Park, and Ram, "Design Science in Information Systems Research", *MIS Quarterly* 28(1), 2004, pp. 75.
- [12] Mendling, J., N. Berente, S. Seidel, and T. Grisold, "Pluralism and pragmatism in the information systems field: the case of research on business processes and organizational routines", *The Data Base for Advances in*

- Information Systems* 52(2), 2021.
- [13] Yamauchi, Y., and T. Hiramoto, “Performative Achievement of Routine Recognizability: An Analysis of Order Taking Routines at Sushi Bars”, *Journal of Management Studies* 57(8), 2020, pp. 1610–1642.
 - [14] Cohen, M.D., and P. Bacdayan, “Organizational Routines Are Stored as Procedural Memory: Evidence from a Laboratory Study”, *Organization Science* 5(4), 1994, pp. 554–568.
 - [15] Feldman, M.S., B.T. Pentland, L. D’Adderio, and N. Lazaric, “Beyond Routines as Things: Introduction to the Special Issue on Routine Dynamics”, *Organization Science* 27(3), 2016, pp. 505–513.
 - [16] Farjoun, M., “Beyond Dualism: Stability and Change As a Duality”, *Academy of Management Review* 35(2), 2010, pp. 202–225.
 - [17] Dittrich, K., S. Guérard, and D. Seidl, “Talking About Routines: The Role of Reflective Talk in Routine Change”, *Organization Science* 27(3), 2016, pp. 678–697.
 - [18] Berente, N., K. Lyytinen, Y. Yoo, and J.L. King, “Routines as Shock Absorbers During Organizational Transformation: Integration, Control, and NASA’s Enterprise Information System”, *Organization Science* 27(3), 2016, pp. 551–572.
 - [19] Østerlund, C., K. Crowston, and C. Jackson, “Building an Apparatus: Refractive, Reflective, and Diffractive Readings of Trace Data”, *Journal of the Association for Information Systems* 21(1), 2020.
 - [20] Dittrich, K., “Ethnography and Routine Dynamics”, In *Cambridge Handbook of Routine Dynamics*. Cambridge University Press, Cambridge, 2021.
 - [21] van der Aalst, W.M.P., *Process Mining: Discovery, Conformance and Enhancement of Business Processes*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
 - [22] Dumas, M., M. La Rosa, J. Mendling, and H.A. Reijers, *Fundamentals of business process management*, Springer Berlin Heidelberg, New York, NY, 2018.
 - [23] Reinkemeyer, L., ed., *Process Mining in Action: Principles, Use Cases and Outlook*, Springer International Publishing, Cham, 2020.
 - [24] Augusto, A., R. Conforti, M. Dumas, et al., “Automated Discovery of Process Models from Event Logs: Review and Benchmark”, *IEEE Transactions on Knowledge and Data Engineering* 31(4), 2019, pp. 686–705.
 - [25] Gaskin, J., N. Berente, K. Lyytinen, and Youngjin Yoo, “Toward Generalizable Sociomaterial Inquiry: A Computational Approach for Zooming in and Out of Sociomaterial Routines”, *MIS Quarterly* 38(3), 2014, pp. 849–A12.
 - [26] Bose, R.P.J.C., W.M.P. van der Aalst, I. Žliobaitė, and M. Pechenizkiy, “Handling Concept Drift in Process Mining”, In R. King, ed., *Active Flow and Combustion Control 2018*. Springer International Publishing, Cham, 2011, 391–405.
 - [27] Maaradji, A., M. Dumas, M.L. Rosa, and A. Ostovar, “Detecting Sudden and Gradual Drifts in Business Processes from Execution Traces”, *IEEE Transactions on Knowledge and Data Engineering* 29(10), 2017, pp. 2140–2154.
 - [28] Carmona, J., B. van Dongen, A. Solti, and M. Weidlich, *Conformance Checking: Relating Processes and Models*, Springer International Publishing, Cham, 2018.
 - [29] Berti, A., S.J. van Zelst, and W.M. van der Aalst, “Process Mining for Python (PM4Py): Bridging the Gap Between Process-and Data Science”, 2019.
 - [30] Vaishnavi, V.K., and W. Kuechler, *Design science research methods and patterns: innovating information and communication technology*, Crc Press, 2015.
 - [31] Hevner, A.R., “A three cycle view of design science research”, *Scandinavian Journal of Information Systems* 19(2), 2007, pp. 4.
 - [32] Gregor, S., L. Chandra Kruse, and S. Seidel, “The Anatomy of a Design Principle”, *Journal of the Association for Information Systems forthcoming*, 2020.
 - [33] Parmigiani, A., and J. Howard-Grenville, “Routines Revisited: Exploring the Capabilities and Practice Perspectives”, *The Academy of Management Annals* 5(1), 2011, pp. 413–453.
 - [34] Nunamaker, J.F., M. Chen, and T.D.M. Purdin, “Systems Development in Information Systems Research”, *Journal of Management Information Systems* 7(3), 1990, pp. 89–106.
 - [35] Pentland, B.T., J. Recker, and I. Kim, “Capturing Reality in Flight? Empirical Tools for Strong Process Theory”, *ICIS 2017 Proceedings*, 2017.
 - [36] van Zelst, S.J., B.F. van Dongen, and W.M.P. van der Aalst, “Event stream-based process discovery using abstract representations”, *Knowledge and Information Systems* 54(2), 2018, pp. 407–435.
 - [37] van der Aalst, W., A. Adriansyah, A.K.A. de Medeiros, et al., “Process Mining Manifesto”, *International Conference on Business Process Management*, Springer Berlin Heidelberg (2011), 169–194.
 - [38] Pentland, B.T., and M.S. Feldman, “Organizational routines as a unit of analysis”, *Industrial and Corporate Change* 14(5), 2005, pp. 793–815.
 - [39] Pentland, B.T., and M.S. Feldman, “13 Issues in empirical field studies of organizational routines1”, *Handbook of organizational routines*, 2008, pp. 281.
 - [40] Feldman, M., “Making process visible: Alternatives to boxes and arrows”, *The SAGE handbook of process organization studies*, 2017, pp. 625–635.
 - [41] Pentland, B.T., J. Recker, and G. Wyner, “Rediscovering Handoffs”, *Academy of Management Discoveries* 3(3), 2017, pp. 284–301.
 - [42] Kremser, W., and G. Schreyögg, “The Dynamics of Interrelated Routines: Introducing the Cluster Level”, *Organization Science* 27(3), 2016, pp. 698–721.
 - [43] Wurm, B., T. Grisold, J. Mendling, and J. vom Brocke, “Business Process Management and Routine Dynamics”, In *Cambridge Handbook of Routine Dynamics*. Cambridge University Press, Cambridge, 2021.
 - [44] Augusto, A., J. Mendling, M. Vidgof, and B. Wurm, “The Connection between Process Complexity of Event Sequences and Models discovered by Process Mining”, *arXiv:2106.07990 [cs]*, 2021.