Enterprise Modelling and Information Systems Architectures
Vol. 15, No. 5 (2020). DOI:10.18417/emisa.15.5

The MobIS-Challenge 2019

A Report on the WI-2019-Workshop on Model-Based Compliance in Information Systems

Stephan Baier[a], Sebastian Dunzer[a], Peter Fettke[b], Constantin Houy[b], Martin Matzner[a], Peter Pfeiffer[b], Jana-Rebecca Rehse[c,b], Martin Scheid[*,b], Sebastian Stephan[b], Matthias Stierle[a], Brian Willems[b]

[a] Chair of Digital Industrial Service Systems, Friedrich-Alexander-Universität Erlangen-Nürnberg, Fürther Str. 248, 90429 Nürnberg, Germany
[b] Institute for Information Systems (IWi) at the German Research Center for Artificial Intelligence (DFKI) and Saarland University, Campus, Building D3.2, 66123 Saarbrücken, Germany
[c] Junior Professorship of Management Analytics, University of Mannheim, L15 1–6, 68161 Mannheim, Germany

Abstract. *Information systems (IS) can significantly support the organization of business processes. However, the proceeding digitalization of processes can also lead to an increasing organizational complexity and the need to more intensely investigate the adherence to external or internal compliance rules. Process-related data from IS and underlying process models can, however, also contribute to an effective compliance checking. This paper summarizes the motivation, the setup, the data set and the results of the 2019 MobIS-Challenge which was conducted as a workshop at WI 2019 in Siegen, Germany. Results submitted to the challenge are presented in detail and directions for future work are discussed.*

Keywords. GRC • Governance, Risk and Compliance • BPM • Business Process Management • Process Mining • Data Set • DCR • Declarative • Dynamic Condition Response Graph

Communicated by Agnes Koschmider. Received 2020-07-24. Accepted on 2020-08-17.

1 Introduction

Nowadays, business processes are increasingly carried out digitally with the help of information systems (IS), which can significantly support an effective, efficient and flexible management of business processes. However, the proceeding digitalization of business processes can also lead to an increasing complexity of organizations and the need to more intensely analyze and ensure the adherence to externally or internally specified compliance rules. In this context, process data from IS and underlying business processes as well as business process models can, however, also considerably contribute to an effective compliance checking. The usage of data from IS allows for an easier reconstruction of business processes, e. g. based on built-in logging mechanisms, and furthermore facilitates the identification of violations of internal or external compliance rules. In this context, there are some research streams, such as process discovery and conformance checking, that develop new methods and techniques to analyze the process data logged by information systems and to use the gained insights for the benefit of the company.

* Corresponding author.
E-mail. martin.scheid@dfki.de
Note: Parts of this article have already been published in: Houy C. et al. (2019) Model-Based Compliance in Information Systems – Foundations, Case Description and Data Set of the MobIS-Challenge for Students and Doctoral Candidates. In: Proceedings of Internationale Tagung Wirtschaftsinformatik 2019 (WI-2019), February 24th-27th, Siegen, Germany, pp. 2026–2039

This report presents the motivation, set up, data set and the results of the WI 2019-Workshop "MobIS-Challenge for Students and Doctoral Candidates: Model-Based Compliance in Information Systems", the topic's conceptual foundations and the use case which was investigated by the challenge participants. Participants of the MobIS-Challenge were supposed to identify opportunities to use IT tools (e. g. existing process mining tools, BPM solutions, self-developed programs, etc.) for analyzing and improving the business process compliance in the addressed business travel management scenario, specifically by examining the compliance of this process and pointing out its weaknesses.

The remainder of this paper is organized as follows: in Sect. 2, we introduce some conceptual foundations of model-based compliance management with a particular focus on business process data from information systems. In Sect. 3, we describe the travel management use case to be analyzed in more detail, providing a verbal description as well as an according business process model, which served as a basis for the simulation and development of the data set. Sect. 4 provides a detailed description of the data set, focusing on the most severe compliance violations included in the approximately 6,500 business travel management cases contained in the data. Sect. 5 presents the tasks which had to be addressed in the challenge as well as a brief description of potential solutions by the authors of this paper. Sect. 6 describes the solution by Dunzer, Baier, Stierle and Matzner in detail before Sect. 7 thoroughly reports on the solution by Willems and Pfeiffer. Sect. 8 concludes the paper.

2 Business Process Compliance

Business process models can serve as an instrument to express and clarify the course of activities in the context of value creation in organizations (Curtis et al. 1992). While business processes can be understood as sequences of executions for the purpose of creating goods and services (Scheer 1999), business process models are representations of business processes which provide the basis for several different tasks of Business Process Management (BPM) (van der Aalst 2013), such as process implementation, execution, controlling or systematic process improvement (Houy et al. 2010).

In order to support their daily operations, business organizations use information systems (IS), like enterprise systems (ES) for enterprise resource planning (ERP), supply chain management (SCM), or customer relationship management (CRM). Such IS – no matter whether they are process-oriented and explicitly produce so-called process log data or not – generate data, which can serve to obtain a view of the underlying business processes. The data generated by IS can, thus, also serve for the identification of compliance violations.

In literature, compliance is one major aspect of the comprehensive topic addressed by the umbrella term Governance, Risk, and Compliance (GRC). GRC and its related policies and rules, as well as technical support approaches and methods, are supposed to ensure a good, responsible and sustainable management of organizations, which follow the applicable law and commonly accepted standards (Becht et al. 2003; Schäfer et al. 2011). Compliance management is supposed to ensure the conformity of "business processes, operations and practice [...] with a prescribed and/or agreed set of norms" (Sadiq and Governatori 2010). In this context, external and internal compliance requirements can be differentiated. Typical external compliance requirements are legal initiatives like the Sarbanes-Oxley Act (SarbOx) in the US or Basel III in the financial sector as well as the so-called Bribery Act 2010 as an anti-corruption legislation example passed in the UK. Furthermore, many organizations have defined internal compliance requirements, which have not been formulated by external authorities, but which are supposed to ensure voluntary conformity of the organization's behavior with common standards.

The term business process compliance is ambiguous and used to address different concepts

Table 1: Dimensions of Compliance Checking (Fellmann and Zasada 2014)

Nr.	Dimension	Sub-Dimensions
1	Scope	Order and Occurrence, Information, Resource, Time, Location
2	Lifecycle phase	Design, Execution, After Execution
3	Formality	Verification / Validation, Business-oriented
4	Contribution type	Technical Artefact, Method, Other

in literature. Some contributions aim at checking the compliance of business process instances (as-is processes) in terms of a defined process model, focusing on their identity without looking at operational business issues. Other contributions refer to business process compliance as a means of checking the operational compliance of an organization based on the underlying business processes, e. g. using process logs to identify and investigate potential violations of external or internal compliance rules (Delfmann and Hübers 2015; Morana et al. 2014; Schultz 2013, 2015; Seeliger et al. 2016). In this contribution, we use the term business process compliance according to the latter understanding.

Fellmann and Zasada (2014) give a comprehensive overview of the current state-of-the-art in their review contribution investigating a total of 84 business process compliance approaches. They identified different dimensions for compliance checking (p. 5) which are described in Tab. 1.

Known approaches for a business process-oriented compliance checking, e. g. use process mining techniques (Caron et al. 2013) or so-called control patterns (Schäfer et al. 2013). Furthermore, there are several commercial tools, which already implement compliance checking techniques in different contexts (Kochanowski et al. 2014). The following section describes the underlying business travel management case treated in the MobIS-Challenge 2019.

3 Case: Business Travel Management Process

3.1 Case Description and Process Model

The case that we provide for the workshop describes a business travel management process in a medium-sized software consulting company. While the data itself was generated by simulation, the process and its governing compliance rules are inspired by one we have encountered in a recent research project.

The goal of the business travel management process is to keep track of all business trips the employees take and their related expenses, such that they can be invoiced to the respective customer, for whose project the respective trip is taken. As is usual in consulting companies, the employees travel quite often to meet with customers, but as software consulting includes some work that can be done remotely, they are not constantly traveling. In order to better control the bookings, to take advantage of economies of scale in the booking process, and to avoid lengthy reimbursements, the company has decided to install a separate travel department, where multiple travel agents are responsible for booking business trips, always in accordance with the respective employee.

To improve the internal process organization, the company has developed its own internal workflow management system which can be accessed by each employee. Travel management is fully covered and logged by this system, with the travel management process implemented as a workflow and the tasks and rights assigned according to the employee's role in the company.

Within the process, there are four acting roles:

1. the Employee, who wants to go on a business trip,
2. the Manager, who has to approve the trip and the expense report,
3. the Travel Department, which is responsible for bookings and price information, and
4. the Accounting, which is responsible for calculating and reimbursing costs.

In our case company, there are 300 employees, 15 managers (including 3 directors), 5 travel agents, and 10 accounting clerks.

The process starts when an employee files a travel request. Such a request offers two options: The employee can either directly file a request or initiate a preliminary price inquiry, which is helpful if the travel expenses and potential booking options are unclear. In this case, the request is forwarded to the travel department, where a travel agent provides a booking proposal and discusses it with the employee. If the employee accepts this proposal, she can adapt her price inquiry accordingly and then transform it into an official travel request. If the proposal is not accepted, she first has to check whether the trip is still necessary, and all data is up-to-date before requesting an update of the booking proposal from the travel agent.

When the travel request is officially filed, it has to be checked for approval before it can be handled by the travel department. The workflow system automatically checks whether the request fulfills the approval requirements and forwards it to the responsible manager for approval. The manager checks the request and either approves it, rejects it, or asks for a correction. In the latter case, the system redirects the request to the employee, such that he can make adjustments according to the manager's requests. This correction process is repeated until the manager finally approves or rejects the travel request.

Once a travel request is approved, or if approval is not required, it is forwarded to the travel department and assigned to a travel agent, who checks whether the request requires any bookings. If not, for example if the employee takes her own car or a company car for a business trip, there is nothing left to do for the travel department. If yes, the travel agent prepares a booking proposal according to the employee's specifications and sends it back for approval. If the employee approves the proposal, the travel agent confirms and pays for all bookings (e. g. hotel, flights, or rental cars). If the employee does not approve the proposal, she has to confirm the data and relevance of the travel request, before she can ask the travel department for an updated booking proposal.

After a business trip is concluded, the employee has to fill out a travel expense report in order to be reimbursed for any travel-related costs. To ensure correct accounting procedures, employees also have to fill out a report if no expenses have incurred. Therefore, the employee first needs to check whether she has any travel-related expense documents (e. g. invoices or receipts). If such documents exist, they have to be uploaded in a digital form, such as a scan. Afterwards, the employee fills out the travel expense report as provided by the workflow system. The confirmed report is automatically forwarded to the respective manager for approval. If the manager decides that the report cannot be approved, it is sent back to the employee for corrections. After the report is approved, the accounting department is in charge of calculating the total travel costs, archiving the travel-related documents, and paying the expenses of the employee.

3.2 Compliance Rules

Within the travel management process, there are a number of external and internal compliance rules, which must be followed. The internal rules are mainly important to keep the accounts correct and up-to-date, whereas the external rules are necessary for invoicing travel expenses to the customer. In detail, the travel management process conforms to the following compliance rules:

1. For each business trip, an according travel request must be filed and, if necessary, approved before the beginning of the trip.

2. Business trips must be necessary to ensure the success of a project. If this necessity cannot be documented, a manager might reject the request.
3. A travel request must contain realistic cost estimations. The real costs in the travel expense report filed after the trip should not exceed this estimation.
4. If the estimated travel cost does not exceed 500€, it does not need to be approved.
5. If the estimated travel cost exceeds 500€, the trip must be approved by the employee's responsible manager.
6. Managers' trips must be approved by a director. The three directors approve each other's trips.
7. Managers should promptly approve, reject, or react to incoming requests.
8. If possible, trips should be booked and paid for by the travel department.
9. After a trip has ended, the travel expense report should be filed immediately.
10. All travel-related expenses should be documented with a receipt.
11. Travel expense reports must be approved by the employee's responsible manager. Again, managers' travel expense reports are approved by a director and the directors approve each other's reports.
12. Managers should check requests and expense reports carefully and ask for corrections, if they find any rule violations.

4 Data Set

4.1 Model Development

In order to generate a viable process log for the MobIS-Challenge case, Scheid et al. developed a process model that we could use as the basis for data generation.[1] We used the ARIS simulation component for data generation and modelled our process as an Event-Driven Process Chain. ARIS is a software tool for the management, execution, and analysis of models like Business Process Models. After the simulation, we enriched it with data that is required for simulation. According to the description above, we separated the process into two major parts. Large parts of the travel planning process are shown in Fig. 1. While the price inquiry handling is described on the left side of the complete process (not displayed in the excerpt in Fig. 1), the right side of the process model deals with approving travel requests in the loop on the left and handling the bookings in the subprocess on the right. The second part of the process, expense reporting, is shown in Fig. 2. First, the employee uploads all travel-related documents and produces an expense report, which is then approved by the manager in the loop on the bottom left. Expenses are reimbursed by the accounting department on the bottom right.

ARIS offers a multitude of attributes to be defined for each model element, some of which were necessary to ensure that our simulation would produce a viable data set. In our case, we needed to define an executing role for each function to assign resources in the process log, the number of employees that instantiate each role, probabilities for each XOR-connector to determine the path frequency, execution times for each function to allocate sufficient time for its execution, and schedules for employees to account for normal working hours. We also associated the start event with an instantiation schedule and a fluctuating delay to introduce some randomness into the start times of each case.

As can be seen in the process model, each function is associated with a role that is responsible for its execution; roles without an explicitly associated role are automatically executed by the workflow system itself. Each role is associated with a schedule, which determines its typical working hours. Employees can work anytime between 6 a.m. and 11 p.m., which factors in that consultants sometimes do organizational tasks like travel management at odd hours. The administrative personnel, i. e. the travel agents and accounting clerks, work typical office hours from 9 a.m. until

[1] Data set source: Scheid, M., Rehse, J.-R., Houy, C., & Fettke, P. (2018). Data Set for MobIS Challenge 2019 [Data set]. https://doi.org/10.13140/RG.2.2.11870.28487.

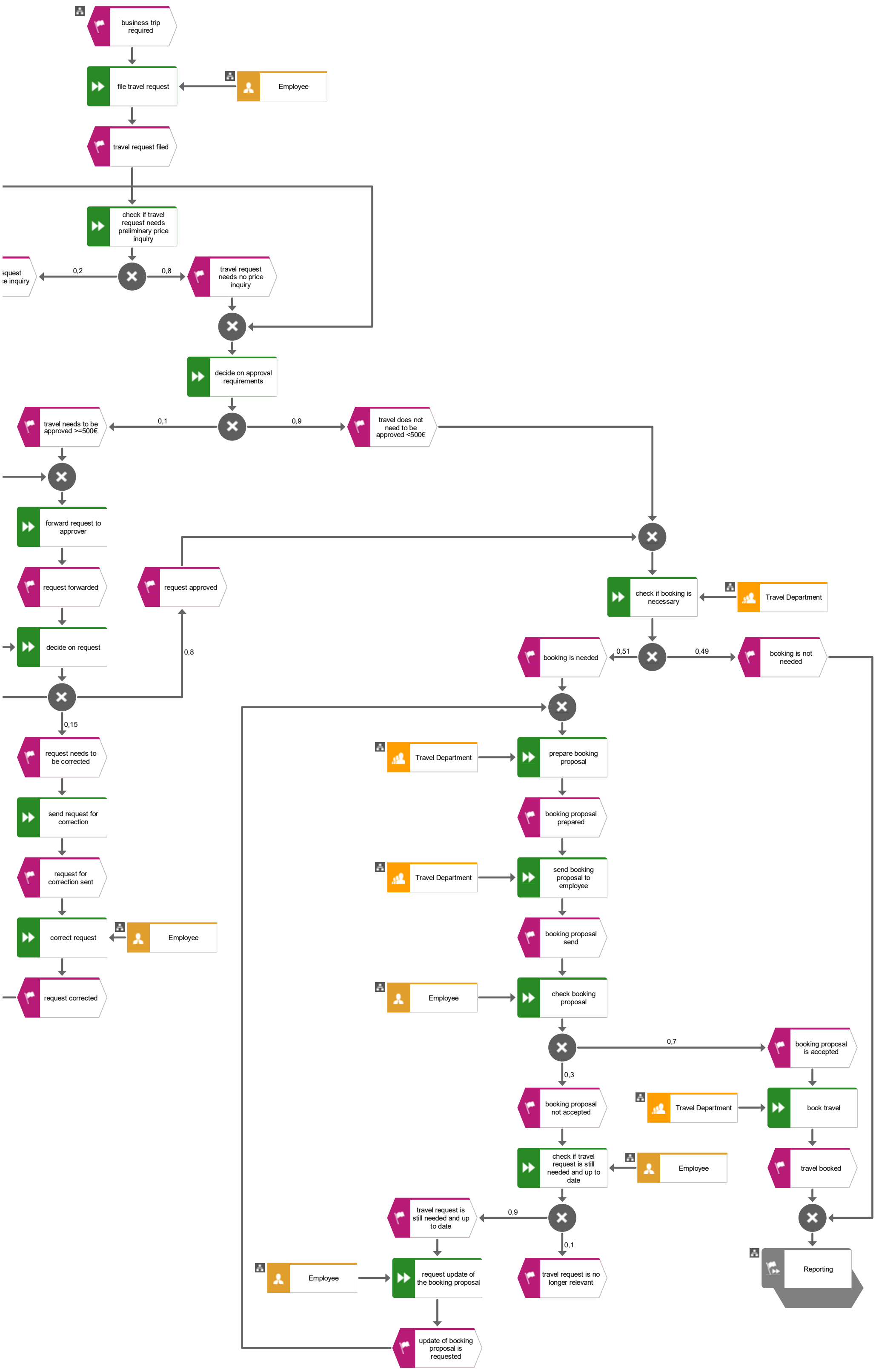

Figure 1: Excerpt of the planning part of the travel management process

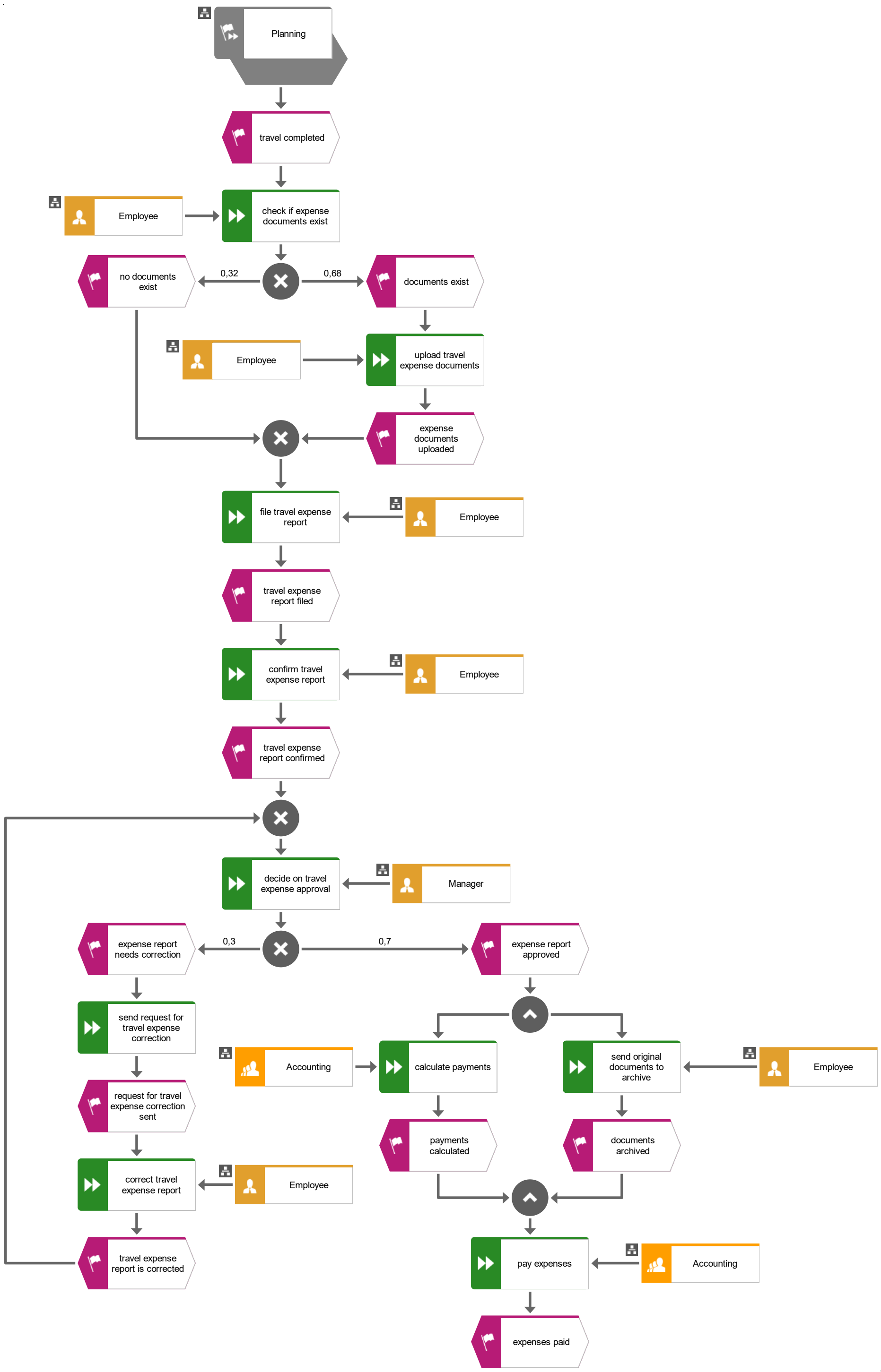

Figure 2: The reporting part of the travel management process

5 p.m. We included one week of vacation time in July, where no travel agent was working. In the managers' schedule, we defined their working hours to be only between 1 p.m. and 2 p.m. to account for the fact that managers are typically very busy and only tend to administrative tasks like travel management at certain hours, e. g. after lunch.

We also defined a fluctuating execution time for each function. ARIS allows several options to set either strict or varying execution times, and we decided to model them all as a capped normal distribution, specifying the expected value, standard deviation, minimum and maximum time for each function. Realistic values were set for each function, such that automated functions executed by the workflow system only take a few seconds, whereas time-intensive functions, like preparing a booking proposal, take several hours. The number of employees (300) and managers (15) in the company were predefined before simulation, but the necessary number of travel agents had to be determined empirically, such that some, but not too many cases were piling up at any given time. 5 travel agents turned out to be a sufficient number.

For most functions, we did not have to define static waiting times to account for employees being busy with other tasks that have nothing to do with travel management. All waiting times for the travel department are caused dynamically because other cases are handled first. Only the accounting functions wait statically for a few days to account for other responsibilities of the accounting department.

4.2 Data Generation

After the model was developed, we used it as the basis for simulating the process data. This simulation consisted of multiple steps.

4.2.1 Process Simulation with ARIS

Based on the developed process model, we used the ARIS simulation component to generate execution data. To account for the travel time that occurred between the two process parts, we connected them with an artificial function ("travel"), with fluctuating static waiting times (to account for the time passing between a travel request and the trip itself) and execution times (to account for the duration of the trip). We wanted to simulate data for one year, so the simulation period was set from October 1st 2016 to December 31st 2017, with the first 91 days functioning as a warm-up phase to have plenty of cases in the system. ARIS used the specified process data to simulate its execution. Each simulation took about 15 to 20 minutes to complete. We exported the case data from ARIS and converted it into a '.csv' file to proceed further.

4.2.2 Generating Additional Data with Excel

The ARIS simulation was only able to generate the process steps itself, so we had to enrich the log with additional data on travel costs and the organizational structure. We defined the company's internal organizational structure, such that we could assign each case to an employee and the responsible manager. There were three cost values that had to be generated, the estimated travel costs, the real travel costs, and the reimbursed costs. The estimated costs were calculated randomly, depending on the length of the trip and whether or not the travel request has to be approved. The real travel cost was calculated to randomly fluctuate in both directions around the estimated cost. Finally, the reimbursed costs depend on whether the travel department booked the trip for the employee. If yes, they were lower than the travel costs; if no, the two numbers were equal.

4.2.3 Manual Data Cleaning

After all data was generated, we had to manually go over it to remove some mistakes and irregularities, such as business trips during Christmas time. After cleaning the data and introducing compliance violations (explained in the next section), our final data set contains 6,555 cases with 26 activities and a total of 83,256 events.

4.3 Violations of Compliance Rules

After the process log was simulated, enriched, and cleaned, we introduced compliance violations.

Table 2: Compliance violations caused by simulation parameters

ID	Compliance violation	Simulation cause
1	Long delays in manager's reaction	Managers are encouraged to answer promptly to incoming travel requests, to avoid not reacting in time for the trip. (Rule 7)
2	Long delays in expense reporting	Accounting should promptly reimburse the employees for their expenses, to avoid unnecessary payment legacy. (Rule 9)
3	Real travel expenses significantly exceed calculated expenses	Employees should give a realistic estimation of expected travel costs, to ensure necessary approvals and facilitate accounting. (Rule 3)

From a data perspective, there are two types of compliance violations, those that were caused by simulation parameters and were already present in the log and those that had to be entered manually.

4.3.1 Violations caused by Simulation Parameters

We defined our simulation parameters (schedules, waiting times, cost calculations) such that we deliberately build some compliance violations directly into the log. They are listed in Tab. 2 and shortly explained in the following.

Violation (1) stems from a time restriction. According to their time plan, managers spend one hour each day for administrative tasks, causing requests to build up. This artificial restriction causes a bottleneck in the process, such that travel requests are delayed for several days before being approved or declined. Violation (2) can be accounted to fluctuating static waiting times, which we attributed to the functions to simulate other responsibilities of the accounting department. Finally, violation (3) is due to the fact that during cost simulation, we had real travel expenses fluctuate in relation to calculated expenses, such that they sometimes are much higher.

4.3.2 Manually entered Compliance Violations

However, most compliance violations (especially those that deviated from the normal process flow) could not be built directly into the log but had to be manually entered. These violations, which include 9 of the 12 in total, are listed in Tab. 3. For each violation, we explain how it contradicts our compliance rules and give its frequency in the log, i. e. the number of existing cases which we altered to violate compliance in the described way.

The following section gives a brief description of the tasks and potential solutions to the problem to be treated in the MobIS-Challenge 2019. Furthermore, sections 6 and 7 present the solutions discussed at the MobIS-Challenge 2019.

5 Challenge Tasks

The participants of the MobIS-Challenge were supposed to identify opportunities to use IT tools (e. g. existing process mining tools, BPM solutions, but also self-developed programs) for analyzing and improving process compliance and pointing out process weaknesses, even beyond conformance issues.

The tasks and leading questions for the MobIS-Challenge were as follows:

1. Describe the process depicted in the log with graphical means. From this description, derive meaningful compliance rules that go beyond the [.] description [given in the call for papers].
2. Which compliance violations can be found in the data? How can these be prevented?
3. Beyond compliance issues, which weaknesses in the process or the organization can be found in the data? How could these be improved?

Table 3: Manually entered compliance violations

ID	Compliance violation	Compliance explanation	Freq.
4	Travel request is submitted after the trip	Travel request must be filed and approved before the trip. (Rule 1)	8
5	Only price enquiry is submitted	Price enquiry must be converted into a travel request before the trip. (Rule 1)	12
6	Manager approves his own trip	Manager's trips must be approved by a director (four eyes principle). (Rule 6)	5
7	Manager approves his own expense calculation	Manager's expense calculations must be approved by a director (four eyes principle). (Rule 6)	10
8	Trip is approved by the wrong manager	Trips must be approved by the employee's own manager. (Rule 5)	11
9	Employee travels despite rejected travel request	Travel requests must be approved by a manager to ensure their necessity. (Rule 1+2)	2
10	Multiple travel requests (less than 500€) for one trip	Requests only have to be approved if the estimated costs exceed 500€. (Rule 4+5)	8
11	New travel request after rejection	The manager assessed the trip as unnecessary for the project success. (Rule 2)	3
12	Paid expenses exceed calculated expenses	The accounting department paid more to the employee than costs incurred for the trip. (Rule 3)	17

4. Which additional insights can be drawn from the data? You can use any tool to develop interesting additional insights in a creative way.

In the next chapters, the results of the winners of the MobIS-Challenge 2019 will be presented. The participants use a wide range of technologies from Dynamic Condition Response Graphs or Neural Networks and show how those technologies can be used to identify compliance violations in the provided log files.

6 Conformance Checking with Dynamic Condition Response (DCR) Graphs — An application to the MobIS Challenge 2019 by Dunzer, Baier, Stierle and Matzner

To address the task of the MoBIS-Challenge 2019, we developed a conformance checking technique that bases on a declarative process modelling language. A systematic literature review of conformance checking identified existing techniques and outlined their key differences (Dunzer et al. 2019). Results indicate an absence of methods for conformance checking in two dimensions. First, the majority of existing techniques focus on the control-flow only and neglect further perspectives like roles or time. Given the relevance of context within process execution and the complexity of business process compliance in organizations, we argue for taking a multi-perspective view in conformance checking. Second, conformance checking techniques are mostly utilizing procedural process models, whereas only a few articles use declarative process models. While procedural languages are comprehensible for well-structured processes, they often lack flexibility. In contrast to procedural languages, declarative languages are useful to express ill-structured processes with a high degree of variation regarding circumstantial information (Fahland et al. 2009; Leoni et al. 2015). Following the process mining research agenda (van der Aalst and Weijters 2004), we enhanced this technique to consider further perspectives, such as organizational resources and data, apart from the control-flow. The implemented technique takes a DCR Graph as well as an event log in the '.xes' format as input and delivers a conformance rate and a list of the identified deviations as output.

6.1 DCR Graphs process modelling

Before stepping into the analysis, we provide a brief overview of the process modelling language DCR Graphs based on Mukkamala (2012). We use a web-based tool (www.dcrgraphs.net) to visualize our process models.

As in most process models, there are activities and connections in DCR Graphs as well. DCR Graphs come with overall two types of activities, (default) activities and nested activities. In the latter, we focus on default activities only. A particular difference to other process modeling techniques is that the process model itself is not static. During the run-time of a process, each activity can be of three distinct states *included* (*In*), *executed* (*Ex*) and *pendingResponse* (*Pe*). Note that the state *In* indicates that a process model includes a particular activity at the current state of execution. If a process model does not include an activity, all other states of the activity become irrelevant since it is then considered non-existent. Activities that are not *In* may not happen. The state *Ex* shows whether an activity has been executed at least once at the current point of execution. Last, the state *Pe* forces the execution of the activity before the process instance terminates. To make use of these states in process execution, DCR Graphs comprise six types of connections whereby each of these represents one particular constraint. For our purposes, we do not include the so-called *spawn* connection. Therefore, there are in total the constraints *include* and *exclude*, *milestone*, *response*, and *condition*. If the source node of a constraint is not *In*, not only the source node is non-existent but also all the connections originating at the particular node.

Last, *guards* enhance these connections with data awareness. *Guards* are Boolean expressions that are related to a connection. If the expression evaluates to *true*, the connection is active and must thereby be evaluated. In contrast if the evaluation results in *false*, the constraint is considered as non-existent for the current step of execution (Slaats et al. 2013). Note that *guards* are an optional element of constraints. If no guard is related to a constraint, it is always active.

The above mentioned DCR Graphs specifications result in four types of violations. First, an activity that is not *In* occurs. Second, an activity that is both *Pe* and *In* at the end of the process execution. Third, an activity targeted by a condition constraint happens even though the source activity is not *Ex* before and *In*. Fourth, the target activity of a milestone is executed, although the source activity is *Pe* and *In*. For the full documentation of the process modelling language of DCR Graphs we refer to Mukkamala (2012) since we only provide a basic view for understanding our process models.

6.2 Analysis method

Following Shearer's CRISP-DM framework, we examined the business and the data to gather prerequisite knowledge about the process specifications and environment (Shearer 2000). As the main tools for this preceding analysis of the process log, we utilized Celonis, a process mining tool, and Microsoft Excel. Note that we draw the business constraints underlying our further analysis – presented in Tab. 5 – exclusively from the data. Hence, the business constraints that we identified slightly differ from the original business compliance rules in Tab. 3.

Since our self-developed conformance checking approach expects a '.xes' file as an event log input, we converted the original '.csv' file with ProM. Furthermore, we modelled the constraints in different process models using the specifications of a DCR Graph. Finally, we conduct the conformance checking with the modelled DCR Graphs and present our findings in the next section.

6.3 Results

In this section, we accumulate the constraints into three distinct process models, whereby each model concerns a different angle of the process.

Table 4: Relation types (Mukkamala and Hildebrandt 2010)

Type	Symbol	Semantics
Include	$\alpha \Rightarrow + \beta$	An include relationship creates a relationship between two activities so that the target activity is *In* after the execution of the source activity.
Exclude	$\alpha \Rightarrow \% \beta$	An exclude relationship creates a relationship between two activities, so that the target activity's *In* is removed after the execution of the source activity.
Response	$\alpha \bullet\!\!\Rightarrow \beta$	A response creates a relationship between two activities so that when the source activity is executed the target activity becomes *Pe*.
Condition	$\alpha \Rightarrow\!\bullet \beta$	A condition creates a relationship between two activities so that the target activity can only occur if the source activity is already *Ex* or not *In*.
Milestone	$\alpha \Rightarrow \diamond \beta$	A milestone creates a relationship between two activities so that the target activity and the source activity can occur. However, if the source activity becomes *Pe*, the target activity cannot be executed until the source activity is *Ex*.

Table 5: Business constraints: Travel expense process

No.	Business constraints
C_1	All expenditures of more than *250 €* must be approved by a *'Manager'*.
C_2	When the activity *'send request for travel expense correction'* occurs, both the *'correct travel expense report'* and the *'decide on travel expense approval'* activities must occur.
C_3	If the activities *'correct travel expense report'* or *decide on travel expense approval* are pending for execution, *'pay expenses'* may not be performed.
C_4	Every case must have a value for *cost* for each occurrence of the activity *'file travel request'*.
C_5	Every case must have a *start and end date* for each occurrence of the activity *'file travel request'*.
C_6	If the activity *'file travel request'* occurs, *'decide on travel expense approval'* must be executed before the payment of the expenses.
C_7	The costs in *'pay expenses'* should not exceed the cost estimate in *'file travel request'*.
C_8	If the costs in *'pay expenses'* exceeds 100 €, the activity *'calculate payments'* must have been executed before.
C_9	The costs in *'calculate payments'* must be similar to the actual expenditure in *'pay expenses'*.

For the reason of comprehensibility, we dispense with merging these three process models since we focus on the results of our conformance checking rather than on a single complex process model.

After analyzing the context of the process and the event log, we elaborated a first DCR Graph, shown in Fig. 3, that comprises the constraints C_1 to C_3. This process model focuses on the case of only paying the expenses after receiving approval. The *Accounting* department can pay expenses without limitations when the price is below 250 € (our assumption), which is visualized by the condition. Otherwise, the *Manager* has to give his approval first. The blue response connection deals with the travel-expense-report correction. In this case, the *Manager's* approval and the *Employee's* correction is mandatory. Last, once a correction is requested, and the approval and correction become *Pe*, expenses cannot be paid until both of these activities occurred. The purple milestone induces this particular constraint.

By performing conformance checking using our developed technique with this process model, we found 178 process instances that violated the constraints, which results in a non-conformance rate of 2.27%. Approximately 95% of the non-conforming traces failed since *Decide on travel expense approval* was absent before a payment

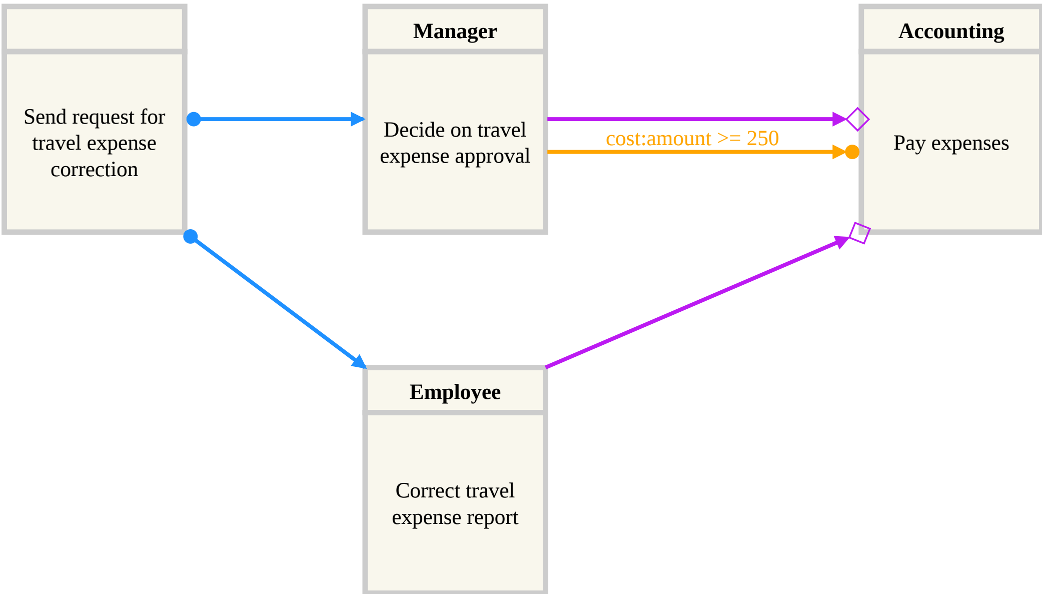

Figure 3: DCR Graph concerned with compliant approval

higher than 250 €. Hence, there are many cases where the payment is made without any approval even though we relaxed the requirement to be only compulsory when the total cost is more than 250 €. Partially this behavior can be explained by missing data, i. e., incomplete cases. Still, we see non-conforming process paths (see Tab. 6) that indicate the completeness of the data and therefore present a valid finding. We identified potential fraudulent cases that impose a financial risk for the company. Such behavior should be restricted in the future.

Table 6: Top three deviating process paths with number of occurrences using the DCR Graph in Figure 3

# of Violations	Process path
90	Pay expenses
38	Send original docs to archive ⇒ Calculate payments ⇒ Pay expenses
37	Calculate payments ⇒ Pay expenses

After examining the travel approval, we created a DCR Graph concerned with the correct form of a filed travel request and filed a travel expense report. Thus, Fig. 4 comprises all constraints related to filing a compliant travel request and a travel expense report. A positive result of the analysis is that all filed travel requests contain a cost estimate. However, in 78 cases, both start and end dates of the travel are missing, which might induce fraudulent behaviour. Especially, the Manager's approval is missing in 1,521 cases. We are well aware that this aspect might be caused by the fact that the instance did not reach the end of the process *'pay expenses'*. But, if we consider this, there are still 1,121 violating cases left representing roughly a sixth of all cases.

Last, we take the compliance of the correct payment into the focus of our analysis. All in all, violating payments lead to a conformance rate of only *29%*. In absolute numbers, 4,629 cases failed. First, 137 cases caused violations due to payments of more than 100 €, without calculating payments beforehand. Second, the cost value in *'calculate payment'* does not equal the actual expenditure in *'pay expenses'*, whereby 518 occurrences resulted in a higher actual expenditure and the remaining 1,108 traces ended in a lower expenditure than calculated. Third, in 1,214 cases, the *Employee's* cost estimate in *'file travel expense'* was exceeded.

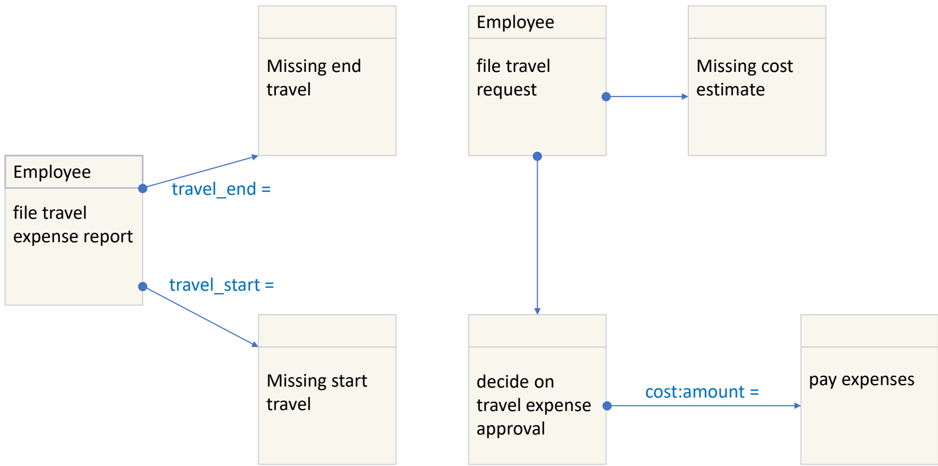

Figure 4: DCR Graph concerned with the correct handling of a request and report

Additionally, we encountered the fact that in 1,217 cases, there is not a filed travel request for the paid expenditure. Therefore, we added the condition to the DCR Graph shown in Fig. 5.

In the present analysis, we examine the given event log in three distinct analysis subjects using a multi-perspective conformance checking approach. Our technique includes both the control-flow of the process and contextual data like costs. Thus, we set up constraints to find compliance problems in the event log following the CRISP-DM approach. We grouped these constraints into the three aspects, (1) managers' approval, (2) correct filing of travel requests and expense reports, (3) compliant payment of expenses. Regarding the aspect of managers' approval, we identify 169 cases where the manager did not approve the trip even though the expenses were higher than 250 €. One result from the analysis is that in 90 of these violating traces, only the activity *'pay expenses'* occurred. In 37 cases, the *Accounting* department calculated the payments before their expenditure. However, the absence of original documents and a travel request might cause incomprehensibility in the future because the IS does not capture the information. Due to the lack of information about the traveling employee and the approval of the business trip's purpose in the event log, these cases are likely to contain non-compliant behavior. Another reason for these process paths might be an undesired workaround where process participants try to avoid bureaucratic activities and delay, whereby they simply skip previous activities. Our proposal to address this issue is that the *'pay expenses'* activities require a related filed travel request. The payment is only released when an employee filed a travel request in the work-flow management system.

The correct filing of travel requests is a minor issue of the subject of analysis. In 78 cases, the travel request neither contained a start nor an end date of a trip. Incomplete data might cause these violations since the business travel might be in an early planning stage. Nevertheless, a solution for this problem is an extra attribute *'estimated date of travel'* in the activity *'file travel request'*. Furthermore, the MobIS-Challenge process description states that a respective manager must approve every travel request regarding its necessity. In the present event log, the requesting employees did not obtain their manager's decision on approval in 1,521 cases. This problem is rather similar, as

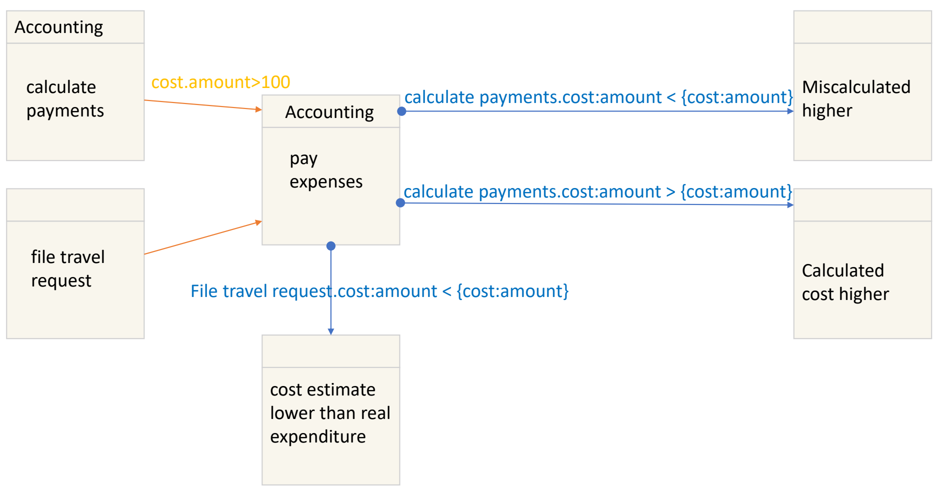

Figure 5: DCR Graph concerned with the correct handling of the payment

the mere payment of the expenses. We recommend blocking the release of the payment until the manager approves the trip.

Last, we identified several problems related to the compliant execution of the payment of expenses. A minor issue might be missing the calculation of costs before its payment in 137 cases. This aspect might be the result of an invoice that already accumulated all expenses related to a business trip or the responsible accountant calculated the costs without the support of IS. Further, in 1,214 cases, the real expenditure exceeded the cost estimate in the filed travel request. Since the company asks the employees to be as realistic as possible, we expected to find occurrences of this violation. Still, this aspect might be caused by employees lowering their cost estimate to increase the likelihood of approval by their manager. To prevent this issue, the company could grant a fixed budget to the traveling employee. After the employee exceeded the budget to a certain amount (e. g., by 10%), the manager reviews the checks and re-approves the expenses. A vital compliance problem we found is caused by the differences in calculated expenses and the actual expenditure. In the overall 1,626 cases, the actual payment exceeded the calculation in 508 cases and was lower in 1,108 cases. Therefore, we strongly recommend introducing the four-eyes principle in the *'Accounting'* department because in more than 4,500 cases, the same accountant calculated the expenses, and paid them afterward. Otherwise, the usefulness of *'calculate payment'* is questionable when the payment differs in roughly 17

The approach we chose bears two limitations, which prevented us from more analyses. First, our developed tool cannot cope with deviating numbers. This aspect would have been particularly useful to examine the differences between the cost estimate and the actual expenses. Then, we could have calculated the compliance using a margin of e. g., 10% maximum deviation between estimate and reality. Second, our algorithm cannot deal with date and time operations until now. Using this feature would have allowed us to investigate whether deadlines of sending specific documents or filing a travel expense report were met.

Despite these limitations, we provided valuable insights into some compliance issues regarding the missing of approvals and the miscalculation of payments.

7 Detection of Compliance Rule Violation in Business Processes using Sequence-to-Sequence Autoencoder by Willems and Pfeiffer

7.1 Introduction

Identification of compliance rule violations in business processes can be a challenging task that is both time- and cost-consuming. Often rules are deduced by hand from experts to check whether a process is compliant or not. Therefore, these rules must be known in advance and have to be manually modified whenever the process changes or new rule violations are found. In order to reduce the time, cost and manual work it takes to detect, update and implement compliance rules, an automatic approach based on a sequence-to-sequence autoencoder is introduced in this section. First, we searched manually for compliance violations using the descriptive tool Disco and python scripts. Afterwards, we built a sequence-to-sequence autoencoder to detect compliance violations automatically. We used the manually found compliance issues to validate the findings of the autoencoder approach. Afterwards, the results are discussed.

7.2 Data Set Analysis

The data set at hand shows the log of a business travel management process in a medium-sized consulting company. An internal workflow management system automatically records the events. Importing the log in the Process Mining tool Disco and generating a process map for process discovery, results in the spaghetti model that is presented in Fig. 6. The model is rather complex with 26 activities and lots of edges between them. In order to get a simpler, better readable model, we reduced the amount of paths. To do that, the paths parameter in Disco was set to 0. An excerpt of the resulting process map is shown in Fig. 7 and presents the normal behavior. This process map is simpler to read and the main path of the process can be identified easily. Besides that, the possible decisions and their mean duration are now visible.

Weaknesses in the Process

Apart from the identification of compliance rule violations, one part of the challenge was concerned with the weaknesses in the process. Fig. 7 shows a simplified process map excerpt of the event log with each activity's mean duration. Darker nodes and edges correspond to a higher mean duration and vice versa. In our analysis, the activities "prepare booking proposal", "book travel", "send original documents to archive" and "calculate payments" are immediately standing out. While the first two activities are performed by the travel department, "send original documents to archive" is executed by the employee and "calculate payments" is done by the accounting. This handover of work between different employees and departments could be an obstacle for optimal processing time. In the case of the travel department, this bottleneck could be explained by over strain of the department's employees. Another reason could be the non-standard manual process of looking for hotels in different cities and different means of transport. The activity "send original documents to archive" is another time-intensive task which, because of regulations, offers not much room for improvements. This task could eventually be obsolete with the change from paper to digital documents. Furthermore, the accounting would also benefit from digital documents. By automating the activity "calculate payments" with a system, this bottleneck would be removed and the mean duration significantly decreased. In addition to that, the edges from "book travel" and "check if booking is necessary" into "check if expense documents exist" have by far the highest mean duration, since this edges represents the actual business trip.

7.3 Compliance Rule Violations in the Data Set

Based on the event log and information given in the challenge, the following compliance rules are considered for our analysis:

1. Business trips have to be approved by a manager.

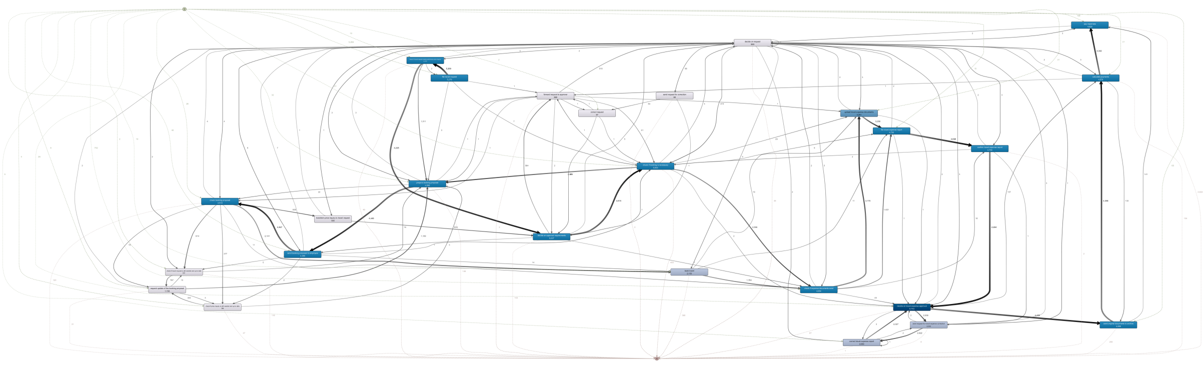

Figure 6: Process map (spaghetti model)

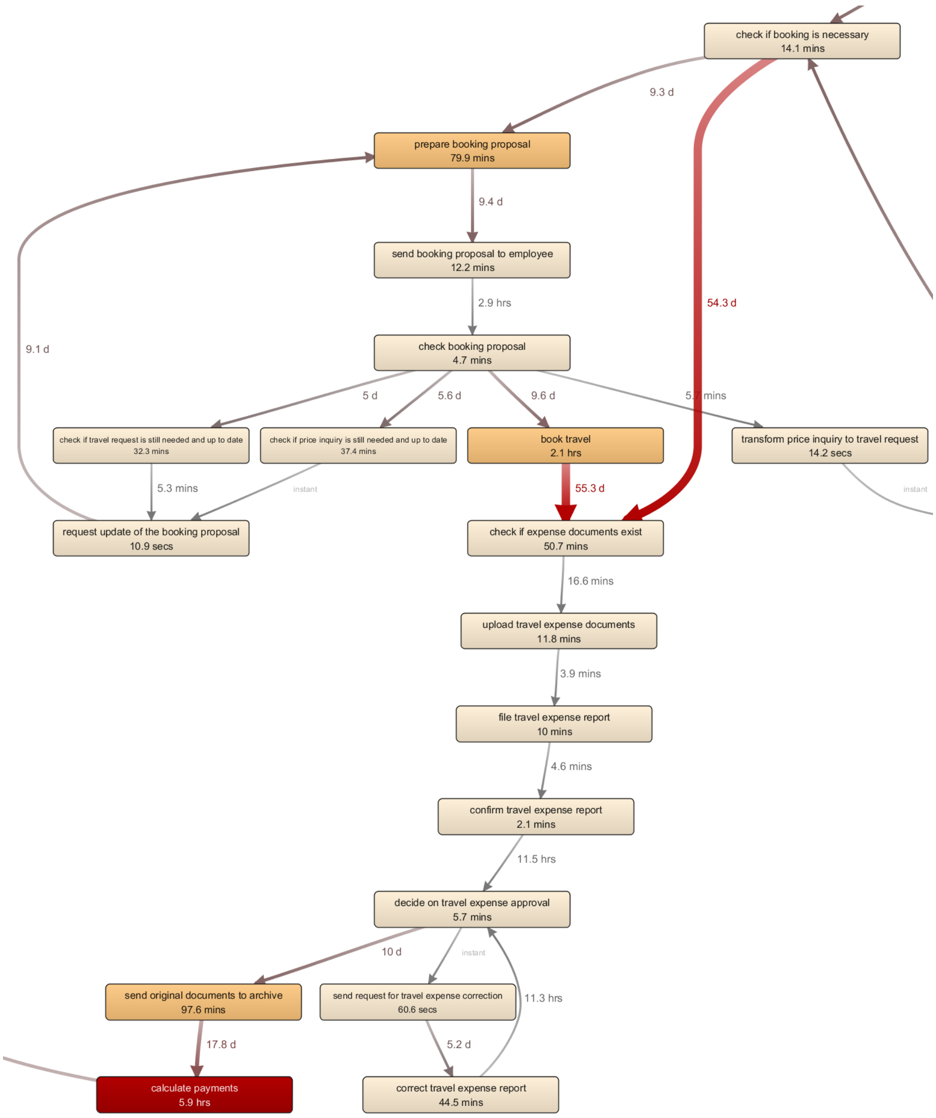

Figure 7: Simplified process map excerpt with total duration

2. Approval has to follow the four-eye-principle.
3. Travel costs have to be estimated as realistic as possible.
4. Travel expenses have to be reported as soon as possible (after the end of a trip).

We further noticed that 1,666 cases did not had any approval by a manager at all. With 333 of that cases were booked without an approval by a manager. Since these 333 cases had costs under 500 it is very likely that business trips under 500 do not require approval by a manager. Therefore, compliance rule 1 could be changed so that all business trips that costs more than 500 have to be approved by a manager.

Rule 1
The first compliance rule to be regarded is the approval of business trips by a manager. To achieve that, the log was filtered with Disco for the activities “decide on request”, “decide on approval requirements” and “decide on travel expense approval” having any role other than manager. Apart from the cases with ID 105 and 3887, the role of every case was left empty. In the two cases mentioned above, the role was specifically set to employee, which implies a violation of this compliance rule while in the other cases we cannot tell exactly who executed the activities.

Rule 2
In order to identify cases where the approval did not follow the four-eye-principle the activities “decide on request”, “decide on approval requirements” and “decide on travel expense approval” were further investigated. Since the initiator of the process has to file the travel request, cases in which the identical resource performed both the activity “file travel request” and one of the three activities mentioned above were searched for. This analysis was achieved by a combination of attribute and follower filter in Disco. The results show that “decide on request” was performed 5 times, “decide on approval requirements” 2 times and “decide on travel expense approval” 7 times by the employee initiating the travel request. In these cases, it was possible for persons to approve their own trips. This could lead to expensive or unnecessary trips which would have been rejected by another person.

Rule 3
As travel expenses have to be estimated beforehand, they are prone to changes throughout the travel process. Whether the estimated costs decrease through better offers or go up because of unforeseen additional expenses. The data set at hand contains three activities in which travel costs are present: “file travel request”, “calculate payments” and “pay expenses”. While the costs regarding “file travel request” are just an estimation, the costs at the other two activities represent the actual and calculated expenses. Nevertheless, high deviations of the costs between these three activities needs to be avoided since it violates the compliance rule of realistic estimations. Therefore, the cost difference between the activities “file travel request” and “calculate payments”, as well as between “calculate payments” and “pay expenses” were analyzed. As a threshold, the difference was set to be at least 1.3 times higher than the previous cost value. Resulting in 29 violations between “file travel request” and “calculate payments”, and 16 violations between “calculate payments” and “pay expenses”.

Rule 4
Since accuracy of the bookings and accounts have to be ensured, the travel expenses need to be reported after the end of the trip as soon as possible. Therefore, the time as days between the end of a travel, based on the according attribute “travel_end”, and the start of the activity “check if expense documents exists” were analyzed. For simplicity reasons, the difference was calculated on a daily level, cutting away any temporal amount below that. The mean time delta is approximately 5 days, with the biggest time difference at 19 days. On the other hand, there are around 150 cases in which the travel expenses got reported in less than a day after the end of the business trip. In order to filter out the outliers which indicate a compliance rule violation the threshold was set to two times the mean. In this data set all cases

which took longer than 10 days to submit the expense documents where extracted, resulting in 246 rule violations. The case lengths of the found violations tend to be longer with a mean of 16.3 in contrast to the mean case length of the other cases at around 13.7.

Prevention of Compliance Rule Violations
Since the whole business travel process is fully covered by an internal workflow management system, the key to prevent compliance rule violations lies within this system.

Rule 1 deals with the approval of the process by a manager. The best way to ensure that solely a manager can approve a request is to develop and implement an access rights concept. This would ensure that the process could only advance if a person with correct rights approves at certain points. Another aspect regarding rule 1 is an improved logging mechanism. As mentioned before, there are just two cases where a role is specified at a certain activity and left empty in all other cases.

In the case of rule 2 it was possible for managers to approve their own travel related decisions. Similar to the recommendation for rule 1, an access rights concept inside the workflow management system would ensure the desired compliance. Furthermore, the organizational structure may need to be updated in order to define who approves the travel process of a manager. Here, interdependencies between the managers have to be organized, to prevent arrangements between them.

Rule 3 is concerned with the realistic estimation of travel expenses. Reducing the deviations of costs during the process cannot be prevented. However, it is possible to gather new information during the process by updating the cost estimation more often. This goes hand in hand with an improved logging strategy which should be examined. Furthermore, there are some cases that contain typos inside the attribute cost. For example, case with ID 304. Here, the costs between the activities "calculate payments" and "pay expenses" increase up from 436.8 to 1436.8. Since the two values differ in just one digit, that strongly indicates a typo and could be checked during the execution.

The main aspect regarding rule 4 is the report of travel expenses as soon as possible. That can be achieved by reminding the employee, after the end of his trip, to submit his expenses. In addition to that, the manager could relieve the employee of some of his work, creating free time for the employee to finalize his travel process.

Apart from the specific recommendations for the single rules, it would be beneficial to have a real time compliance checking function which would inform the manager or the controlling if there are compliance violations during the process.

7.4 Sequence-to-Sequence Autoencoder

As mentioned before, finding compliance rule violation in business processes can be a challenging task that is time- and cost-consuming. In the following, an autoencoder approach is presented. Compliance rule violations are treated as anomalies of the business process recorded in the event log and a sequence-to-sequence autoencoder (S2SA) is used to detect them. These networks can be trained in an unsupervised manner with the input of the model and the target to predict being the same – in this work one case of the event log. No compliance rules need to be defined by hand since the autoencoder learns the normal behaviour, that should not contain compliance issues, from the data. By feeding cases to the trained autoencoder the difference between the input and output case, called reconstruction error, is used to detect anomalies.

A case is defined as a sequence of events recorded in the event log. We assume that most of the cases in the event log reflect normal behavior of the underlying process and only a small portion of entries refer to abnormal behavior which have a chance of being ones with compliance rules violated. Since most of the data is not abnormal, the autoencoder model remembers these cases well and reconstructs them with low error. For anomalous ones, the reconstruction error will be higher since they do not occur frequently in the data. In the following, we evaluate if a sequence-to-sequence autoencoder is able to detect the compliance rule violations found in the previous analysis and if any additional insights into the process can be drawn.

7.4.1 Related Work

Compliance checking is an active research field that gets more attention due to the increasing number of laws and regulations business processes need to comply with (Fellmann and Zasada 2014). Many state-of-the-art approaches use rules deviated from experts by hand to check whether a process is compliant or not (Awad et al. 2015, 2008; Montali et al. 2014; Thullner et al. 2011). For this, the compliance rules need to be known in advance. To the best of our knowledge, (Nolle et al. 2016, 2018) developed so far the only approaches to detect anomalous process executions without any further information about the underlying process. However, the approach is not specifically made for compliance detection but performs very well in detecting anomalous cases. In our approach, we use a similar approach as presented in (Nolle et al. 2016) to detect compliance rule violations of a business process recorded in an event log.

7.4.2 Setup and Architecture

The sequence-to-sequence autoencoder consist of two Long Short-term Memory (Hochreiter and Schmidhuber 1997) (LSTM) layers which have shown to perform well on sequences of business process events. The architecture is shown in Fig. 8. Between the two LSTM layers a bottleneck of fully connected layers is applied. The autoencoder is split in an encoder and decoder. The encoder takes one case as a sequence of encoded events as input and passes them through the LSTM and fully connected layer ending in a last fully connected layer with two neurons that represents the latent space. The decoder takes the two-dimensional latent space as input and tries to reconstruct the original encoded sequence using several fully connected and one LSTM layers. Before passing the sequence generated by the fully connected layer to the LSTM decoder, the sequence is reversed. As mentioned in Shalev-Shwartz and Zhang (2014) this makes it easier for a sequence- to-sequence model to learn. Internal states of the encoder LSTM layer are passed to the decoder LSTM layer without going through the bottleneck.

Before passing the event log to the model it is preprocessed. Since the goal is to detect compliance rules that occur in the sequence of events or attributes, all numerical values are deleted. Columns start, end, travel_start, travel_end and cost are dropped resulting in an event log with 26 different activities, 5 types and 335 different resources. Within these there exist 3,742 unique combinations of these 3 attributes. Each combination is represented by an integer in the range of 0 to 3,742 as encoding. Since the number of events in a case differs, we padded them with zeros to equal length. The encodings are scaled to (0, 1) and split into a training and test set. In front of the first LSTM layer an embedding layer takes the input and transforms it to a higher dimensional representation which is fed into the autoencoder. The weights of the embedding layer are adjusted during training as part of the model. Using an embedding is a frequently used approach in the field of NLP. For the presented approach, instead of representing the combinations of activity, resource and type as one-hot encoding a different encoding using an embedding layer is chosen. This reduced the dimensionality of the input and output and made training easier and faster. Training was performed in epochs on 90% of the data set while 10% were used to evaluate the model reconstruction performance. Between the epochs the data set was shuffled, furthermore we used dropout and early stopping to not overfit to the data set. The reconstruction error is calculated as loss of the mean squared error between the input and output sequence of events.

7.4.3 Results

For the experimental evaluation the same MobIS-Challenge data set as before was used to detect the compliance rule violations discussed in Tab. 7. The model performed well in reconstructing many of the cases with an error close to zero which

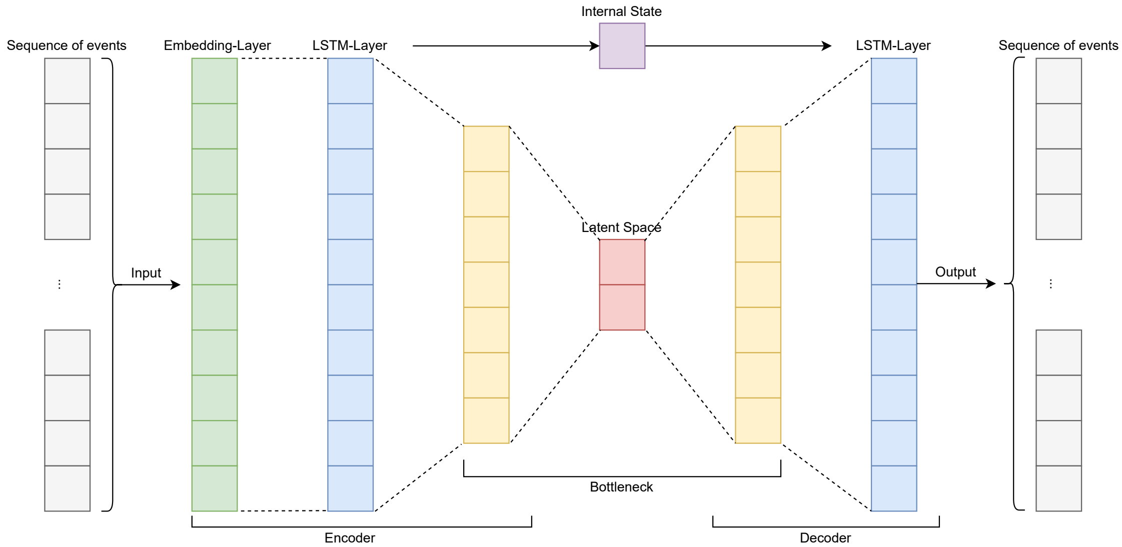

Figure 8: Scheme of the architecture of the sequence-to-sequence autoencoder

Table 7: Manually detected anomalies in comparison with the anomalies found by the sequence-to-sequence autoencoder

Rule	Violations found manually	Violations found by S2SA	Percentage
1	2	0	0%
2	14	7	50%
3	45	10	22,2%
4	246	45	18,29%

shows that it is able to reconstruct the sequence of events as well as the resource and type of each event correctly. For other cases the reconstruction error is still huge. Fig. 9 shows the reconstruction error per event for each case with the annotated rule violations found by the manual analysis. We calculated the error per event since cases differ a lot in their length and therefore in the total sum of loss. Darker reddish color points represent cases with more events.

There are several short cases with high error as well as long cases with low error. As threshold for an anomaly we choose the cases that have a higher reconstruction error per event than 1.5 times the average error per event. This showed to be a good value to neither detect too many nor too few cases.

The number of detected violations is presented in Tab. 7. In total 1,500 cases were found to be anomalous. Within these, the highest total number of found compliance rule violations come from rule 4. 45 out of 246 total violations are detected. This rule is broken if the time between the end of a travel and the submission of the expense report takes longer than 10 days. Since the datetimes are not included in training data the only reference for this rule violation may be the length of the case. The mean length of the detected cases is 7.6 days which is less than the average of all cases as well as the mean of all the cases that violated rule 4.

Different from that, violations of rule 2 can be found by investigating the attributes resource and type across the process. These attributes

Figure 9: Cases with loss per event and marked with violated compliance rule

are part of the training data and therefore the autoencoder should be able to detect them. For our understanding, this kind of rule violation is hard to detect by hand since it resembles a kind of excess of authority that is immanent in the process. With 50% detection rate it seems like the network is indeed able to find these rule violations in several cases.

For rule number 3, the difference in the costs, the detection rate is very low with 18%. Like for rule 4 the costs are not part of the data the model was trained on and therefore this low rate is not surprising. For rule 1, no violations out of 2 are found. With so few occurrences of violation they are pretty rare and difficult to detect.

Apart from the detected compliance rule violations, many cases were found to be anomalous which do not break any compliance rules nor are suspicious. These cases are e. g. abnormal long or short (e. g. feature many loops) or have infrequent attributes. For further analysis, we visualized the latent space as output of the last fully connected layer of the encoder. This layer has only two neurons that act as the bottleneck of the model and are nice to visualize. Fig. 10 shows the latent space. The latent space starts to separate the datapoints into different regions. In contrast to the loss, the datapoints are not grouped by the length of the case. There are some points far outside referred as global outliers. Violations of rule 4 are grouped into four to five different regions with some points outside these groups. Most of the compliance rule violations are still close together and right next to the cases with no violations which makes it difficult to detect them in the latent space.

8 Conclusions

In this contribution, we have presented the results of the WI 2019-Workshop “MobIS-Challenge for Students and Doctoral Candidates: Model-Based Compliance in Information Systems”. We provided a detailed description of the underlying problem, which had to be investigated by the challenge participants. Additionally, we described how the problem and the data was generated and

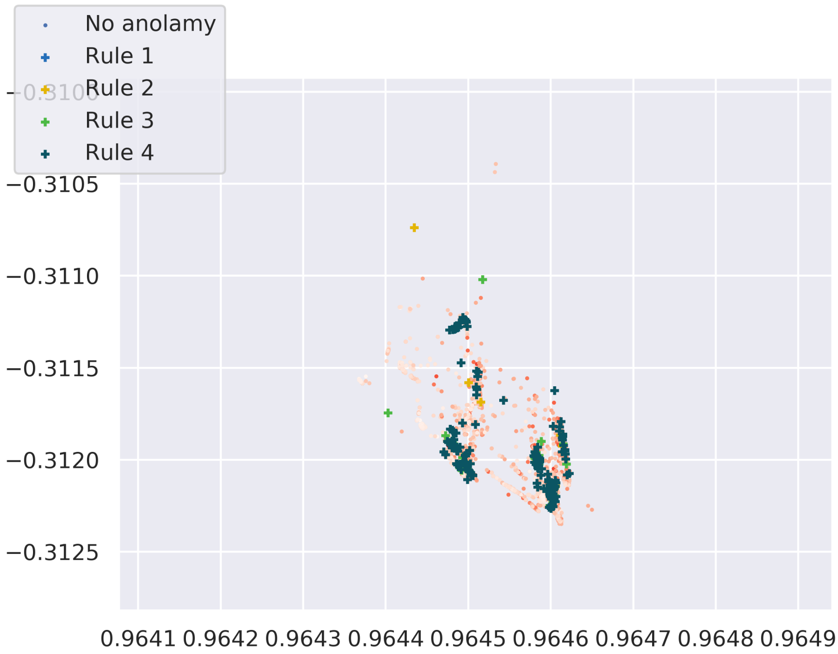

Figure 10: Latent space with marked compliance rule violations. Darker reddish shades reflect latent representations of longer cases

edited to match the challenge requirements. The requirements themselves were taken over from a preliminary project and adapted and prepared for the challenge.

Two submissions for the challenge were accepted, and the participants were invited to the workshop. The participants addressed the MobIS-Challenge using quite different approaches leading to different, but nevertheless very interesting results, which provide different perspectives on the problems to be treated in the challenge. Dunzer, Baier, Stierle, and Matzner implemented and used a process conformance checking technique based on DCR graphs and provided a particular perspective making use of declarative process modeling, which can help to identify specific compliance issues and can, thus, have certain advantages in model-based compliance management.

Willems and Pfeiffer developed a framework for the automated detection of compliance violations based on a sequence-to-sequence LSTM autoencoder and used different state-of-the-art Process Mining tools and frameworks like ProM, Disco, Celonis and bupaR for further analyses of the given event data. Also, their approach provides interesting new insights and showcases the potential of neural network-based techniques in compliance management scenarios. Finally, it became clear that not all compliance validations can be identified by only looking at the structure of the process. Furthermore, a detailed look at the attributes the corresponding distribution and potential anomalies is important. It is also difficult to automatically differentiate between anomalies and compliance violations. A lot more research is needed in this field. The provided data set of the 2019 MobIS-Challenge can serve for and support further BPM research endeavors, e. g. in terms of the validation and evaluation of process mining or data analytics approaches for the investigation of data, which is relevant for business process compliance issues. Hence, BPM and conceptual modelling researchers can use the data set to test and improve their developed methods and techniques.

References

Awad A., Barnawi A., Elgammal A., Elshawi R., Almalaise A., Sakr S. (2015) Runtime Detection of Business Process Compliance Violations: An Approach Based on Anti Patterns. In: Proceedings of the 30th Annual ACM Symposium on Applied Computing. ACM, pp. 1203–1210

Awad A., Decker G., Weske M. (2008) Efficient Compliance Checking Using BPMN-Q and Temporal Logic. In: International Conference on Business Process Management. Springer, pp. 326–341

Becht M., Bolton P., Roell A. (2003) Corporate governance and control In: Handbook of the Economics of Finance Elsevier, pp. 1–109

Caron F., Vanthienen J., Baesens B. (2013) Comprehensive rule-based compliance checking and risk management with process mining. In: Decision Support Systems 54(3), pp. 1357–1369

Curtis B., Kellner M. I., Over J. (1992) Process Modeling. In: Communications of the ACM 35(9), pp. 75–90

Delfmann P., Hübers M. (2015) Towards Supporting Business Process Compliance Checking with compliance Pattern Catalogues - A Financial Industry Case Study. In: Enterprise Modelling and Information Systems Architectures 10(1), pp. 67–88

Dunzer S., Stierle M., Matzner M., Baier S. (2019) Conformance checking: a state-of-the-art literature review. In: Proceedings of the 11th International Conference on Subject-Oriented Business Process Management. ACM, pp. 1–10

Fahland D., Lübke D., Mendling J., Reijers H. A., Weber B., Weidlich M., Zugal S. (2009) Declarative versus imperative process modeling languages: The issue of understandability. In: Enterprise, Business-Process and Information Systems Modeling. Springer, pp. 353–366

Fellmann M., Zasada A. (2014) State-of-the-Art of Business Process Compliance Approaches: A Survey. In: European Conference on Information Systems (ECIS). AIS, pp. 1–17

Hochreiter S., Schmidhuber J. (1997) Long Short-term Memory. In: Neural Computation 9, pp. 1735–1780

Houy C., Fettke P., Loos P. (2010) Empirical research in business process management – Analysis of an emerging field of research. In: Business Process Management Journal 16(4), pp. 619–661

Houy C., Rehse J.-R., Scheid M., Fettke P. (2019) Model-Based Compliance in Information Systems – Foundations, Case Description and Data Set of the MobIS-Challenge for Students and Doctoral Candidates. In: Proceedings of Internationale Tagung Wirtschaftsinformatik 2019 (WI-2019), February 24th-27th, Siegen, Germany, pp. 2026–2039

Kochanowski M., Drawehn J., Kötter F., Renner T. (2014) Compliance in Geschäftsprozessen. Business Process Management Tools 2014. Fraunhofer

de Leoni M., Maggi F. M., van der Aalst W. M. P. (2015) An alignment-based framework to check the conformance of declarative process models and to preprocess event-log data. In: Information Systems 47, pp. 258–277

Montali M., Maggi F. M., Chesani F., Mello P., van der Aalst W. M. P. (2014) Monitoring Business Constraints with the Event Calculus. In: ACM Transactions on Intelligent Systems and Technology 5(1), 17:1–17:30

Morana S., Schacht S., Scherp A., Maedche A. (2014) Designing a Process Guidance System to Support User's Business Process Compliance. In: International Conference on Information Systems (ICIS). AIS

Mukkamala R. R. (2012) A formal model for declarative workflows - dynamic condition response graphs. PhD thesis, IT University of Copenhagen

Mukkamala R. R., Hildebrandt T. T. (2010) From dynamic condition response structures to Büchi automata. In: 2010 4th IEEE International Symposium on Theoretical Aspects of Software Engineering. IEEE, pp. 187–190

Nolle T., Seeliger A., Mühlhäuser M. (2016) Unsupervised Anomaly Detection in Noisy Business Process Event Logs Using Denoising Autoencoders. In: International conference on discovery science. Lecture Notes in Computer Science. Vol. 9956. Springer, pp. 442–456

Nolle T., Seeliger A., Mühlhäuser M. (2018) BINet: Multivariate Business Process Anomaly Detection Using Deep Learning. In: Business Process Management 2018. Lecture Notes in Computer Science. Vol. 11080. Springer, pp. 271–287

Sadiq S., Governatori G. (2010) A Methodological Framework for Aligning Business Processes and Regulatory Compliance. In: Handbook of Business Process Management 2. Strategic Alignment, Governance, People and Culture. Springer, pp. 159–176

Schäfer T., Fettke P., Loos P. (2011) Towards an Integration of GRC and BPM - Requirements Changes Caused by Externallly Induced Complexity Drivers. In: BPM 2011 International Workshops, Part II. Lecture Notes in Business Information Processing. Vol. 100. Springer, pp. 344–355

Schäfer T., Fettke P., Loos P. (2013) Control Patterns - Bridging the Gap between IS Controls and BPM. In: European Conference on Information Systems (ECIS) Vol. Paper 88. AIS

Scheer A.-W. (1999) ARIS – Business Process Frameworks. Springer

Schultz M. (2013) Enriching Process Models for Business Process Compliance Checking in ERP Environments In: Design Science at the Intersection of Physical and Virtual Design, LNCS 7939 Springer, pp. 120–135

Schultz M. (2015) Business Process Compliance from an Audit Perspective. PhD thesis, Universität Hamburg

Seeliger A., Schmidt B., Nolle T., Mühlhäuser M. (2016) Process Compliance Checking using Taint Flow Analysis. In: International Conference on Information Systems (ICIS). AIS, pp. 1–18

Shalev-Shwartz S., Zhang T. (2014) Accelerated Proximal Stochastic Dual Coordinate Ascent for Regularized Loss Minimization. In: Proceedings of the 31st International Conference on Machine Learning. PMLR, pp. 64–72

Shearer C. (2000) The CRISP-DM model: the new blueprint for data mining. In: Journal of Data Warehousing 5(4), pp. 13–22

Slaats T., Mukkamala R. R., Hildebrandt T., Marquard M. (2013) Exformatics Declarative Case Management Workflows as DCR Graphs. In: Business Process Management. Lecture Notes in Computer Science. Vol. 8094. Springer, pp. 339–354

Thullner R., Rozsnyai S., Schiefer J., Obweger H., Suntinger M. (2011) Proactive Business Process Compliance Monitoring with Event-Based Systems. In: 2011 IEEE 15th International Enterprise Distributed Object Computing Conference Workshops. IEEE, pp. 429–437

van der Aalst W. M. P. (2013) Business Process Management: A Comprehensive Survey. In: ISRN Software Engineering, pp. 1–37

van der Aalst W. M. P., Weijters A. (2004) Process mining: A research agenda. In: Computers in Industry 53(3), pp. 231–244