Multi-Dimensional Fraud Detection Metrics in Business Processes and their Application

Badr Omair[1], Ahmad Alturki[2]
Faculty of Computer and Information Sciences
King Saud University, Riyadh, Kingdom of Saudi Arabia

***Abstract*—Occupational fraud is defined as the deliberate misuse of one's occupation for personal enrichment. It poses a significant challenge for organizations and governments. Estimates indicate that the funds involved in occupational fraud cases investigated across 125 countries between 2018 and 2019 exceeded US$3.6 billion. Process-based fraud (PBF) is a form of occupational fraud that is perpetrated inside business processes. Business processes underlie the logic of the work that organizations undertake, and they are used to execute an organization's strategies to achieve organizational goals. Business processes should be examined for potential fraud risks to ensure that businesses achieve their objectives. While it is impossible to prevent fraud entirely, it must be detected. However, PBF detection metrics are not well developed at present. They are scattered, unstandardized, not validated, and, in some cases, absent. This study aimed to develop a comprehensive PBF detection metric by leveraging and operationalizing a taxonomy of fraud detection metrics for business processes as an underlying theory. 41 PBF detection metrics were deduced from the taxonomy using design science research. To evaluate their utility, the application of the metrics was undertaken using illustrative scenarios, and a real example of the implementation of the metrics was provided. The developed metrics form a complete, classified, validated, and standardized list of PBF detection metrics, which include all the necessary PBF detection dimensions. It is expected that the stakeholders involved in PBF detection will use the metrics established in this work in their practice to increase the effectiveness of the PBF detection process.**

Keywords—Business process fraud; fraud detection; fraud indicators; fraud measures; fraud metrics; PBF; red flags

I. Introduction

Fraud refers to an action that is designed to deceive others. Fraud results in a loss for the victim and gain for the perpetrator [1]. The Association of Certified Fraud Examiners (ACFE)[1] defines occupational fraud as the "use of one's occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization's resources or assets" [2, p. 86]. Organizations and individuals alike can be financially or physically affected by fraud [3].

Fraud can either be internal, when it is committed by someone inside an organization, or external, when it originates from outside an organization [4]. In this research, the focus is on internal or occupational fraud.

Fraud is becoming a globally prevalent threat [5]. It is estimated that the overall loss resulting from 2,504 cases of occupational fraud that were investigated between January 2018 and September 2019 exceeded US$3.6 billion across 125 countries [2]. The ACFE estimates that organizations lose approximately 5% of their revenues to fraud each year [2]. The wave of financial scandals that has been sweeping the world in the current century has also heightened the awareness of the need to manage fraud risk [6].

Process-based fraud (PBF) is a form of fraud that occurs in business processes. It can be identified by measuring the deviation from the process model [7]. However, deviation in the business process model is not always regarded as fraud; in order to confirm that fraud has taken place, a domain expert must investigate the matter.

Business process refers to a collection of related events, activities, decision points, actors, and objects that lead to an outcome that is valuable to at least one customer [8]. Business processes are core assets of organizations [8], and they are essential in the implementation of organizational strategy [9]. Business processes should be examined to detect any associated potential fraud risks that may threaten the achievement of business objectives [10]. However, at present, PBF detection metrics are not well addressed [11]. They are incomplete, overlapping, scattered, and not standardized [11]. Furthermore, the increase in fraud in recent years reflects the persistent nature of the issue [12]. Therefore, as it is impossible to prevent PBF completely, detecting it when it occurs is essentially.

This manuscript aims to develop comprehensive metrics that cover all the components necessary for the effective detection of PBF. The developed metrics will contribute to the effective detection of PBF as they provide a comprehensive, validated, and standardized list of PBF detection metrics.

First, the metrics are deduced from the taxonomy of fraud detection metrics for business processes [13]. The taxonomy serves as the underlying theory using design science research (DSR). The use of this taxonomy provides a complete understanding of PBF detection, coverage of all PBF detection elements, and a checklist of best practices that define PBF detection metrics [13]. Second, an illustrative scenario, as an evaluation method [14], is provided for each of the developed metrics in order to validate their utility. Ultimately, an implementation that uses the process mining technique is proposed to demonstrate the technical application of the metrics.

[1] https://www.acfe.com

The remaining contents of this paper are organized as follows: Section II provides the background of the topic; Section III explains the methodology followed in the current work; Section IV proposes the complete PBF detection metrics; Section V provides a real example of the implementation of the metrics; Section VI shows and discusses the results; and, finally, in Section VII, the conclusions and direction in which the work in this field may progress in the future are presented.

II. Background

Implementing fraud detection and fraud prevention systems is essential for effective fraud risk management [15]. Fraud prevention consists of measures to avoid or reduce fraud. In addition, in fraud detection, measures that help identify fraud when it occurs are used [15]. Since preventing every instance of fraud is impossible, continuous application of fraud detection techniques is necessary to protect against any instances that were not prevented [3].

Fraud detection techniques can be placed into one of three categories [16]. First, the misuse-based detection technique uses a predefined list (i.e., known patterns) of possible fraud schemes to detect fraud. It is an expert fraud detection system that uses predefined metrics. Its advantage is a low false alarm rate, but it cannot detect instances of fraud that follow new patterns [16]. Second, the anomaly-based technique can be implemented using machine learning techniques, which leads to the detection of any suspicious behavior that deviates from standard behavior [17], [18]. It does not require a predefined list of fraud schemes, and it can detect new cases of fraud. However, it suffers from a high false alarm rate [19]. Third, the hybrid technique attempts to combine the previous two techniques to overcome their limitations [16].

Successful fraud detection must include an examination of business processes to identify the potential origins of fraud [20]. Business processes are the core of business process management (BPM), which is a management discipline that uses business processes to implement organizational strategy [9]. It is a management discipline that requires continued focus, and often, significant changes in management style [9].

PBF detection metrics form the intersection between fraud risk management and BPM, as reflected in the bidirectional arrow mentioned in Fig. 1. The use of such metrics is common in fraud risk assessment and process monitoring and control[2] which are elements of fraud risk management and BPM, respectively.

Fraud detection can be achieved using a taxonomy to predefine initial fraud schemes [1], [21]. A taxonomy is a set of dimensions, each consisting of a set of mutually exclusive and collectively exhaustive characteristics [22]. A taxonomy of fraud detection metrics for business processes was proposed in [13], as depicted in Fig. 2. The taxonomy provides a holistic view of fraud detection in business processes. It consists of the dimensions examined in the following subsections.

A. Fraud Domain

This dimension covers the application domain of fraud detection. Knowing the fraud domain is crucial in the detection of fraud because it allows an understanding to be gained of the problem domain [23]. In addition, specific fraud, which is particular to certain domains, exists, and these cases require special handling. This dimension contains two characteristics:

- *General*: Describes all metrics that can be used in any application domain.
- *Specific:* Covers a particular application domain, such as finance.

B. Fraud Data Scheme

This dimension covers all the potential fraud schemes in the data. Fraud data schemes provide a list of possible data schemes used for committing fraud, which means that understanding them is critical for detection. This dimension contains the following data schemes:

- *Anomalous:* Covers any data that can be characterized as ambiguous or exceptional (e.g., too long, too short, excessive, and outliers).
- *Discrepant:* Describes inconsistent data (e.g., the conflict between input and output, and between past and current).
- *Missing:* Covers insufficient and absent data.
- *Wrong:* Covers incorrect data (e.g., inaccurate, non-conforming, fictitious, error, and outdated).

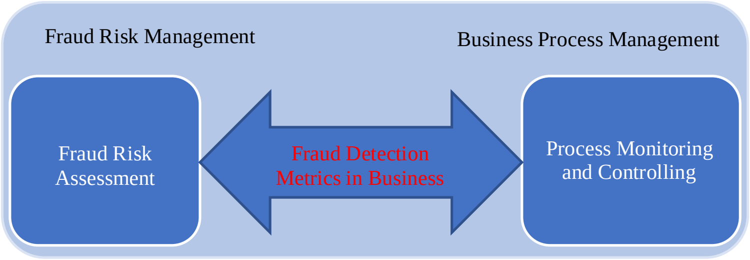

Fraud Risk Management
Business Process Management
Fraud Risk Assessment
Fraud Detection Metrics in Business
Process Monitoring and Controlling

Fig. 1. Execution Scope of PBF Detection Metrics.

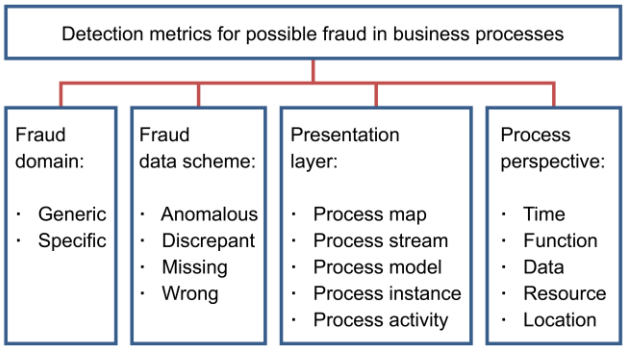


Fig. 2. Taxonomy of Fraud Detection metrics in Business Processes. Source: [13].

[2] Performance measures are usually identified during the process analysis phase of BPM. In some cases, they are identified during the process identification phase [8]. Moreover, business process measures can be classified as measures for business process models and execution [63]. Since fraud detection is the goal, this study focuses on measures executed in the process monitoring and control phase to determine how well the executed processes work with regard to the chosen measures [9].

C. Presentation Layer

This dimension aims to examine all layers of the business processes, as illustrated in Fig. 3.[3] The layers are essential for detecting fraud because every layer can give specific auditing information [24]. Additionally, some fraud cases do not become apparent by looking at a single layer. The dimension contains the following characteristic layers:

- *Process map:* Gives an overview of all business processes and determines their relationships. The process map also contains aggregated data on all business processes in the organization. It is useful for planning fraud detection in business processes.
- *Process stream:* Offers a greater level of detail compared to the process map. It helps set the scope by focusing on a collection of processes that form a specific (and usually vital) business cycle, such as the purchase-to-pay cycle. This layer allows fraud examiners to aggregate data on a particular business cycle.
- *Process model:* Represents a single business process, such as the payment process. It provides more detail on the structure of the process, its controls, activities, and actors. This layer contains aggregated data on many instances of a specific business process.
- *Process instance:* Depicts the details of one particular instance of a process model. It contains concrete data on one specific business process instance, such as payment instance number 123.
- *Process activity:* This is the lowest layer in the *presentation layer* dimension. It can be considered an element of the *process instance* layer with a particular focus. It gives concrete data with more detail on a specific activity in a particular process instance, such as approval activity.

D. Process Perspective

This dimension looks at business process from various angles because, for successful fraud detection, it is necessary to examine all aspects of business process [20]. This dimension contains the following characteristics:

- *Time:* This perspective regards business process's time (e.g., throughput time, actual processing time, waiting time, and deadlines).
- *Function:* This perspective is concerned with the implementation of the activities in business process (e.g., work frequency, work sequence, work decision, process steps, and process control flow).
- *Data:* The data perspective covers all the data that are entered, consumed, and delivered by business process (e.g., process objects).
- *Resource:* This perspective involves all the actors that interact with business process, including customers, software, business role, business units, suppliers, and employees.
- *Location:* This perspective is concerned with the location of business process's execution.

The results of the literature review on PBF detection metrics[4] are summarized in Fig. 4 in the form of a literature map [11]. The literature map illustrates the topics relevant to fraud detection metrics in business processes, as well as the frequency of their recurrence in the literature. Omair and Alturki [11] demonstrated that, at present, the explicitly defined PBF detection metrics, which are listed in Table I, do not adequately address the essential conceptual perspectives of business process.

Combined metrics and process mining can improve fraud detection [25]. Process mining is a methodology that aims to discover, monitor, and improve real processes by analyzing their event logs [26]. It connects model-based process analysis (e.g., simulation) and data-oriented analysis techniques (e.g., data mining) [27]. Process mining associates the actual processes with their data and the process models [28].

Process mining has been successfully applied to detect fraud [23], [29]–[31]. It can reveal fraudulent transactions that cannot be detected using traditional audit methods [29], [32], [33]. Relying on measurements of throughput processing (not just measurements of the input-output relation), process mining can identify a problem's root cause. This involves identifying the process model, and, subsequently, the performance of the process [34].

Using process mining to detect fraud has many advantages. Since event logs are automatically logged in most existing systems [35], it is possible to save time and effort, and to improve detection accuracy by taking real and complete data as opposed to samples [36]. Also, reading from event logs ensures independence from human intervention, which guarantees unaltered and error-free data [37]. According to the ACFE report [2], the median time for detecting fraud is 14 months. During the interval between occurrence and detection, the most significant financial losses tend to occur. However, using online process mining solutions can change this reality [38].

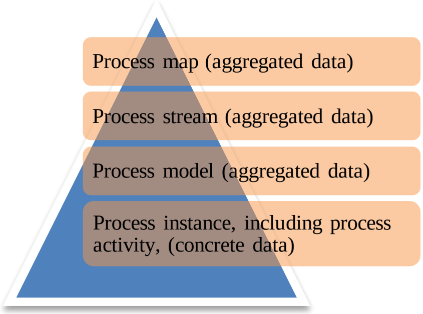

Process map (aggregated data)
Process stream (aggregated data)
Process model (aggregated data)
Process instance, including process activity, (concrete data)

Fig. 3. Presentation Layers of Business Processes.

[3] For more information, see [13], [24].

[4] For the complete literature review and analysis, see [11].

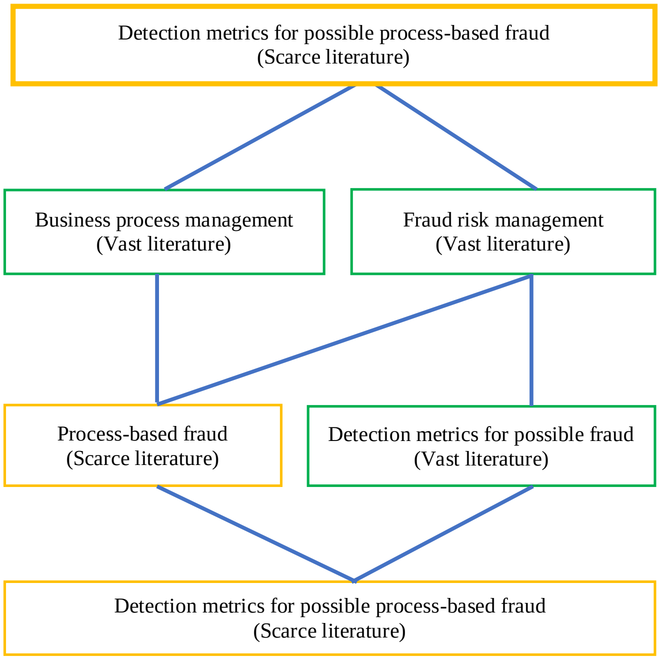

Detection metrics for possible process-based fraud
(Scarce literature)
Business process management
(Vast literature)
Fraud risk management
(Vast literature)
Process-based fraud
(Scarce literature)
Detection metrics for possible fraud
(Vast literature)
Detection metrics for possible process-based fraud
(Scarce literature)

Fig. 4. Literature Map. Source: [11].

TABLE I. FRAUD DETECTION METRICS IN BUSINESS PROCESSES. ADAPTED FROM [24]

ID	Metric name	Explanation	Reference
1	Skipped activity	Not executing an activity that is prescribed in the standard operating procedure (SOP). The skipped activity is either a routine activity or a decision activity [42].	[7], [30], [31], [42]–[45]
2	Wrong resources	The activity is performed by an actor who is not defined in the SOP.	[7], [30], [31], [42]–[46]
3	Wrong duty	The same actor executes different activities, which should require different privileges. This includes “wrong duty sequence” in the sequence activity, “wrong duty decision” in the decision activity, and “combined wrong duty”, a combination of wrong duty sequence and wrong decision sequence [42].	[7], [30], [31], [42]–[45]
4	Wrong pattern	Deviation from the standard sequence prescribed in the SOP.	[7], [30], [31], [42]–[46]
5	Wrong decision	Decision activity execution is a deviation from standard decision execution, as stated in the SOP.	[7], [30], [31], [42]–[45]
6	Wrong throughput time	The activity execution time deviates from the standard time, as stated in the SOP. It includes “wrong throughput time min” and “wrong throughput time max” [42].	[7], [30], [31], [42]–[45]
7	Parallel event	Nonparallel events are performed simultaneously.	[7], [30]
8	Originator behavior	The actor's behavior while executing the activity is anomalous.	[7], [30], [31]

Process mining anomaly techniques include control flow analysis, role resource analysis, throughput time analysis, and decision point analysis [39]. The study undertaken by [4], which proposed a process mining method for PBF detection, suggested the concept “1 + 5 + 1”, which includes (1) log preparation; (5) (a) log analysis, (b) performance analysis, (c) social analysis, (d) conformance analysis, (e) process analysis; and (1) refocusing and iteration. A combination of the red flag approach (i.e., metrics approach) and process mining were proposed in [25] to reduce the false positive rate in detecting fraud. The method connects the red flag approach with process mining by using the red flag to present unusual behavior, whereas process mining involves visualizing the business process flow. In [40], a validated method, based on the most accepted lifecycle model for the implementation of the process mining project [41], was proposed for an application in auditing information systems. It used process mining as an expert system engine to address the limitations of other auditing methods involved in fraud detection, including sampling, due to questionable effectiveness as they lack automation and have a narrow scope.

III. METHODOLOGY

In her remarkable and exceptional work, Gregor [47] explained information systems (IS) theories in terms of five types: analytic theory, explaining theory, prediction theory, explaining and prediction theory, and design and action theory[5] Taxonomy is a taxonomic theory and can be classified as an analysis theory [47]. Analysis theories define or classify specific dimensions or characteristics of individuals, groups, situations, or events by describing the shared features found in discrete observations [47]. These theories answer *what* questions, and they are used as a foundation for developing more advanced theories, as shown in Fig. 5 [47], [48].

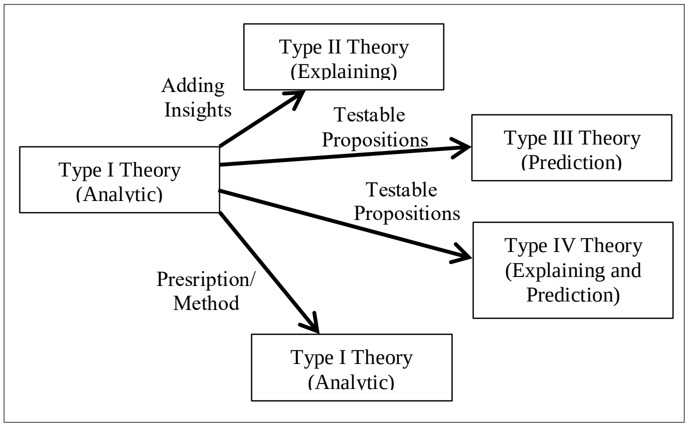


Fig. 5. Evolution of Analytic Theories into other Types. Source: [48].

The DSR methodology can be used to conduct research when the desired goal is an artifact or a recommendation [49]. DSR artifacts are classified into constructs, models, methods, and instantiations [50]. The developed PBF detection metrics are subsumed under the *method* artifact type [51]. This study aims to design an artificial (i.e., human-made) artifact (i.e., PBF detection metric), which fits well within the DSR

[5] For more information, see [47].

paradigm [52]. Furthermore, the pragmatic viewpoint of DSR, which confirms the inability to separate utility from reality [49], is suitable for the nature of the activity of PBF detection.

Following the DSR paradigm, the taxonomic theory [13] was used in this research as the foundation for deriving PBF detection metrics. The taxonomy [13] was developed using DSR's build/evaluate cycle [52], which led to the definition of the building blocks of PBF detection metrics by implementing the method of Nickerson *et al.* [22]. Since taxonomy can be used as a foundation to produce new knowledge [22], [47], [48], [53], the taxonomy of fraud detection metrics for business processes [13] was used deductively to develop the PBF detection metrics (i.e., the taxonomy's objects). Adapting [54], the following steps were taken to develop the metrics:

- Define the measured entity in the study, namely, business process.
- Specify the attributes of the defined entity (i.e., business process), which are already developed by the taxonomy (i.e., the taxonomy's dimensions and characteristics) [13].
- Define the metrics by matching the attributes of the defined entity.

Theoretical validation of the developed metrics can be achieved through the use of a validated taxonomic theory [13]. In addition, in order to evaluate the utility of every developed metric, an illustrative scenario was used [16]. Lastly, an implementation was provided to explain the metrics technical application.

IV. PBF Detection Metrics

Using the taxonomy of fraud detection metrics for business processes as the underlying theory [13], PBF detection metrics can be derived by matching the characteristics of the taxonomy's dimensions. Selecting the matched characteristics depends on the application domain, project scope, and the case situation. However, general PBF detection metrics can be developed by matching the selected characteristics from the process perspectives, presentation layers, and fraud data schemes dimensions.[6] [7] Table II shows the derived list of PBF detection metrics, including the metric's ID, name, description, and the illustrative scenario. The generally derived PBF detection metrics covered all the dimensions of PBF detection (i.e., full-dimensional metrics), as stated in the taxonomy of fraud detection metrics for business processes [13].

TABLE II. General Samples of PBF Detection Metrics

ID	Metric name	Description	Illustrative scenario
1	Wrong activity time	Indicates whether the process activity's time is incorrect.	The execution time of the approval activity in invoice XYZ is not valid.
2	Wrong instance time	Shows whether the process instance's time is incorrect.	The waiting time between activity A and activity B in an invoiced instance exceeds the allowed time.
3	Wrong stream time	Indicates whether the process stream's time is incorrect.	The waiting time between the raising of invoice XYZ and its payment as processes in the purchase-to-pay stream exceeds the allowed time.
4	Discrepant instance time	Shows whether the process instance's time causes conflict.	The throughput time of an activity is longer than the throughput time of the instance that includes the activity.
5	Discrepant stream time	Reveals whether the process stream's time causes conflict.	The execution time of invoice XYZ and its payment as processes in the purchase-to-pay stream are identical.
6	Anomalous activity time	Indicates whether the process activity's time is abnormal.	The execution of the approval activity in invoice XYZ occurred outside of the working hours.
7	Anomalous instance time	Indicates whether the process instance's time is abnormal.	The throughput time of a payment instance is too short.
8	Anomalous model time	Shows whether the process model's time is abnormal.	The execution time of all payment instances for supplier XYZ are all at 8 P.M.
9	Anomalous stream time	Indicates whether the process stream's time is abnormal.	The waiting time between receiving and inspection as processes in the purchase-to-pay stream is very long.
10	Anomalous map time	Indicates whether the process map's time at the *map layer* is abnormal.	The total execution time of all the organization's processes is too short.
11	Wrong activity function	Indicates whether the process activity's work is incorrect.	The decision was incorrectly made in activity XYZ.
12	Wrong instance function	Shows whether the process instance's work is incorrect.	A payment instance must not be executed because the vendor's work is not yet finished.
13	Wrong stream function	Reveals whether the process stream's work is incorrect.	A payment process was executed before the receiving process in the purchase-to-pay stream.
14	Missing activity function	Indicates whether the necessary process activity's work is missing.	The approval activity in invoice XYZ is missing.
15	Missing instance function	Demonstrates whether the necessary process instance's work is missing.	The inspection instance to show that item XYZ was checked is missing.
16	Anomalous activity function	Indicates whether the process activity's work is unusual.	The decision made in activity XYZ was unexpected.
17	Anomalous instance function	Shows whether the process instance's work is unusual.	Unnecessary activities (i.e., excessive work) are performed in executing an invoice instance.
18	Anomalous model function	Indicates whether the process model's work is unusual.	The number of refund instances of customer XYZ is unusual.
19	Anomalous	Shows whether the	Purchase-to-pay processes

[6] Other metrics can be similarly developed by matching the selected characteristics that should be specified for every project.

[7] The selected characteristic of the fraud domain dimension is *general*. This is because the scope of the developed metrics does not focus on a specific fraud domain.

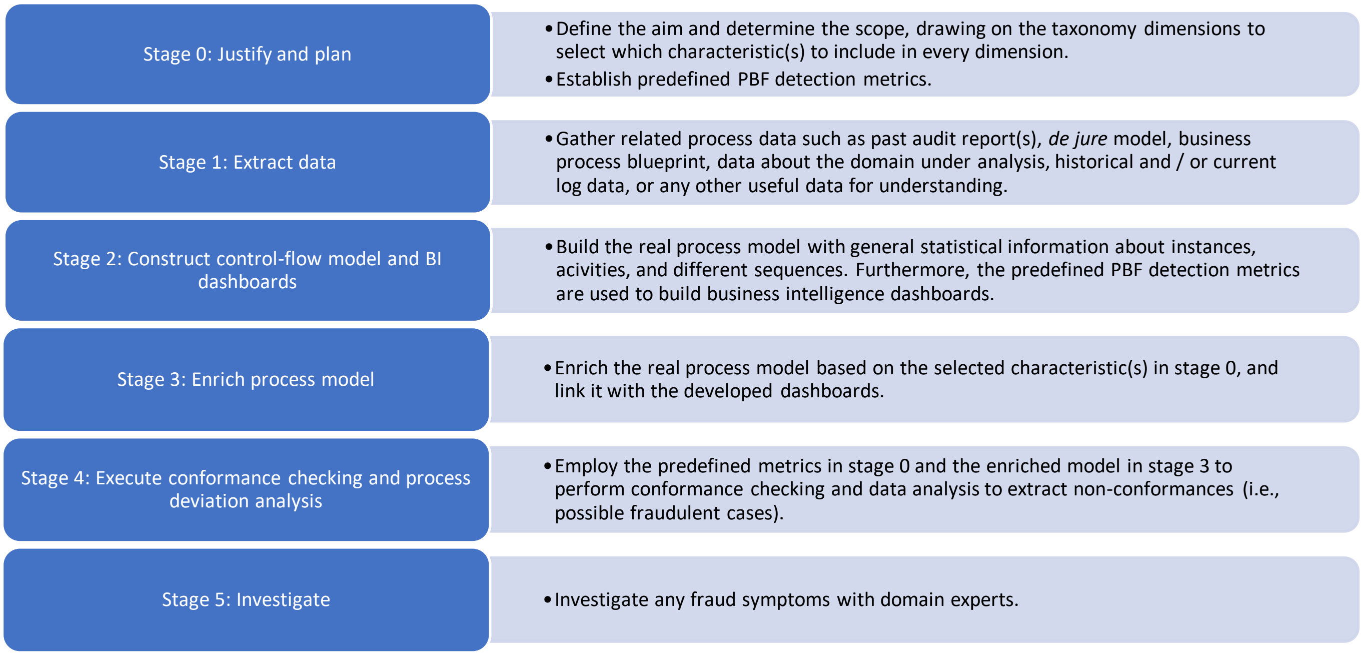
	stream function	process stream's work is unusual.	for supplier XYZ have always had a non-standard process flow without justifications.
20	Anomalous map function	Indicates whether the process map's work at the *map layer* is unusual.	Cancellation of 25% of the organization's processes.
21	Discrepant stream function	Shows if the process stream's work causes conflict.	The payment instances are more than the invoice instances as processes in the order-to-cash stream; however, they should be the same.
22	Wrong activity data	Indicates whether the data produced or consumed by the process activity are incorrect.	The attached document in activity XYZ at invoice A is invalid.
23	Missing activity data	Indicates whether the data produced or consumed by the process activity are missing.	The signature data in activity XYZ of invoice A is missing.
24	Discrepant activity data	Shows whether the data produced or consumed by the process activity are inconsistent.	The attached form in the activity XYZ at invoice A has a signature date that follows the activity date.
25	Discrepant instance data	Shows whether the data produced or consumed by the process instance are inconsistent.	In an invoice instance, the input data of activity B does not match the output data of activity A, though they should be equal.
26	Discrepant stream data	Indicates whether the data produced or consumed by a process stream are inconsistent.	The total amount of orders and the total cash received as processes in the order-to-cash stream should be equal but they differ.
27	Anomalous activity data	Shows whether the data produced or consumed by the process activity are suspicious.	The activity XYZ has unnecessary recorded data (maybe to complicate the auditing process).
28	Anomalous instance data	Indicates whether the data produced or consumed by a process instance are suspicious.	The attached document in the activities A and B of a process instance are in different formats. Even the document should not be different.
29	Anomalous model data	Shows whether the data produced or consumed by the process model are questionable.	The inspection instances of all the items from supplier XYZ always have a lengthy inspection report.
30	Wrong activity resource	Indicates whether the process activity's resource is incorrect.	An employee not authorized to perform activity XYZ in an invoice instance performed it.
31	Wrong instance resource	Shows whether the process instance's resource is incorrect.	Issue and review of invoice XYZ performed by the same employee (violates the separation of duties law).
32	Missing activity resource	Indicates if the process activity's resource is missing.	An anonymous person performed activity XYZ in an invoice instance.
33	Anomalous activity resource	Shows whether the process activity's resource is suspicious.	Activity XYZ in a payment instance, usually executed by employee X is executed by employee Y instead.
34	Anomalous instance resource	Shows whether the process instance's resource is suspicious.	The same employee did most of the activities in a receiving instance XYZ.
35	Anomalous model resource	Indicates whether the process model's resource is suspicious.	The same employee approved all the payments for supplier XYZ.
36	Wrong activity location	Shows whether the process activity's location is incorrect.	The activity XYZ of a receiving inventory instance was executed outside the approved receiving area.
37	Wrong instance location	Reveals whether the process instance's location is incorrect.	Two activities to be executed at the same location for an invoice instance performed at different locations.
38	Anomalous activity location	Indicates whether the process activity's location is suspicious.	The execution location of activity XYZ in a payment instance was very distant.
39	Anomalous model location	Shows whether the process model's location is suspicious.	All the large payments are made only at one location.
40	Anomalous stream location	Indicates whether the process stream's location is suspicious.	Two processes that are usually executed in the same place in the order-to-cash stream were executed at different locations.
41	Missing activity location	Shows whether the process activity's location is missing.	The execution location of activity XYZ in a payment instance is not specified.

V. Implementation

Based on [25], [40], [41], as well as the taxonomy developed in [13], a method can be proposed for implementing PBF detection metrics. The method uses data and process mining to ensure an effective PBF detection process. Both techniques are used to detect fraud in business processes [45], [55]. Although data mining and process mining share many features, the key difference is that data mining aims to discover previously unknown and interesting patterns in the datasets, while process mining focuses on finding process relationships [28]. Thus, data mining techniques for detecting fraud are usually unsuitable for analyzing the behavior of control flow in a business process [39]. However, process mining can be used to assess the control flow of a business process [56] and to analyze process performance, event sequence, and process roles [57]. Still, process mining focuses on the control flow of transactions [56] and not on process content (e.g., transaction value). Therefore, data mining and process mining are both needed.

Real data [58] on purchase-to-pay process events in a multinational paints and coatings company were used for implementation.[8] The implementation method is illustrated in Fig. 6 and described in the following steps:

[8] To reduce data noise, the data were filtered according to document type, item category, and timeframe to include "standard PO," "three-way matching," "invoice before GR," and 2018 (quarter 2).



Stage 0: Justify and plan

- Define the aim and determine the scope, drawing on the taxonomy dimensions to select which characteristic(s) to include in every dimension.
- Establish predefined PBF detection metrics.

Stage 1: Extract data

- Gather related process data such as past audit report(s), *de jure* model, business process blueprint, data about the domain under analysis, historical and / or current log data, or any other useful data for understanding.

Stage 2: Construct control-flow model and BI dashboards

- Build the real process model with general statistical information about instances, acivities, and different sequences. Furthermore, the predefined PBF detection metrics are used to build business intelligence dashboards.

Stage 3: Enrich process model

- Enrich the real process model based on the selected characteristic(s) in stage 0, and link it with the developed dashboards.

Stage 4: Execute conformance checking and process deviation analysis

- Employ the predefined metrics in stage 0 and the enriched model in stage 3 to perform conformance checking and data analysis to extract non-conformances (i.e., possible fraudulent cases).

Stage 5: Investigate

- Investigate any fraud symptoms with domain experts.

Fig. 6. Implementation Steps. Adapted from [25], [40], [41].

Stage 0: At this stage, the scope and aims should be defined after establishing a thorough understanding of the application domain. This includes understanding the business process, identifying the theoretical existence of fraud schemes, cataloging all potential fraud methods and red flags[9], defining the general multi-dimensional metrics by using the taxonomy, and defining specific multi-dimensional metrics for the selected fraud schemes and methods. Every metric may include a metric formula, data source, metric description, data update frequency, metric unit, threshold or compared value, related fraud scheme, and fraud method or red flag.

In this implementation, the aim was to detect fraud in the purchase-to-pay process by examining execution deviations. The scope was determined based on the following dimensions and characteristics of the taxonomy of fraud detection metrics for business processes [13]:

- Fraud domain: In this implementation, the purchase-to-pay business process was selected. Thus, {specific: finance and general} were chosen as the fraud areas for the implementation because general PBF detection metrics are also used.
- Presentation layer(s): {process stream, model, instance, and activity} were selected to satisfy the aim. However, the process stream layer was not included in the implementation due to missing data.
- Process perspective(s): {time, function, data, and resource} were selected. Location perspective data are not available. However, depending on the case situation and data availability, it may be useful to include all process perspectives.
- Fraud data scheme(s): To specify critical data schemes that can effectively detect fraud in this implementation, {anomalous, discrepant, missing, and wrong} were selected. The selection of the fraud data scheme characteristics was based on the case situation and the quality of existing data. However, if possible, it is always useful to include all fraud data schemes.

The selected dimensions, along with their characteristics, ought to assist in developing the predefined metrics. Fraud examiners can also add more useful metrics based on their experience. In this implementation, the generic and specific metrics defined in Appendix A are used based on the case situation and the existing data.[10] The specific multi-dimensional metrics for the fraud schemes and fraud methods are defined based on the common fraud schemes appearing in the fraud tree [10].[11] The fraud tree was selected for the following reasons: (1) it represents a comprehensive classification of the most common occupational financial fraud schemes; and (2) it is developed by a standards organization (ACFE).

Stage 1: At this stage, all the useful process data for detecting PBF should be collected. Examples of data that should be collected are the past audit reports, process events log, and process model, as depicted in Fig. 7 [59]. This model is referred to as the *de jure* model, which represents the desired, ideal, or required process.

[9] Red flags are signs of potentially fraudulent behavior [62].

[10] Sound knowledge of business rules is valuable in defining effective metrics.

[11] For more information about the fraud tree, see [10].

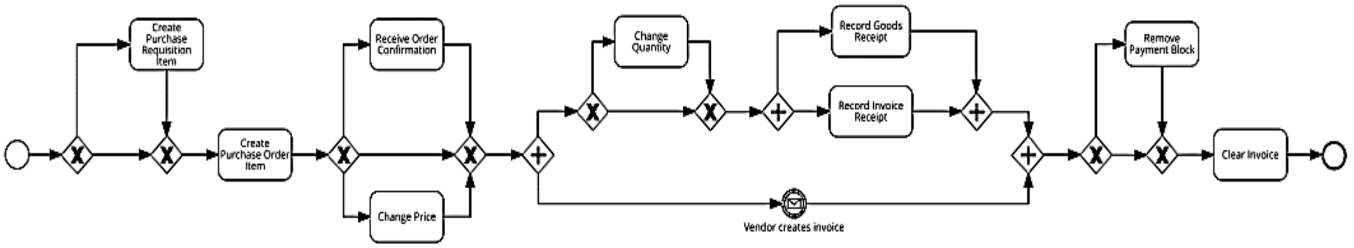


Fig. 7. The De Jure Model for the Purchase-to-Pay Process. Source: [59].

Stage 2: Using the process mining discovery technique, the *de facto* model with general statistical information was constructed as shown in Fig. 8. The *de facto* model describes reality with potential violations [60]. It was implemented using the Celonis process mining software.[12] It is possible for the auditor to analyze differences between the *de jure* and *de facto* models in order to detect fraud [33].

Moreover, the predefined metrics were represented on dashboards, as shown in Appendix B. In this case, the Celonis process mining software was also used.

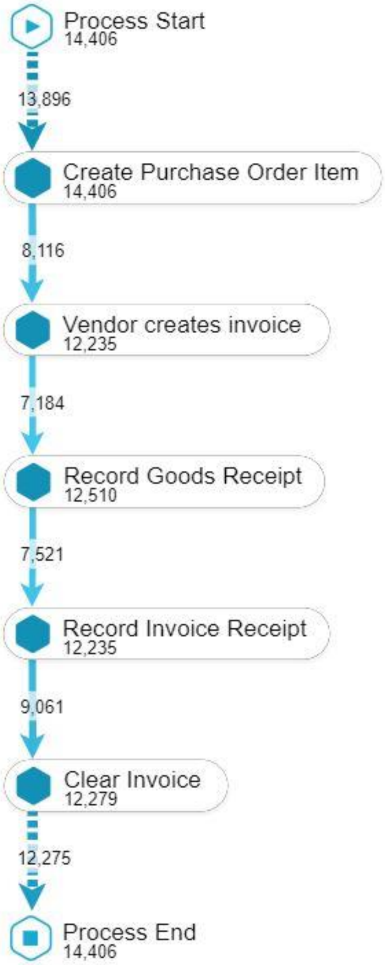


Fig. 8. The De Facto Model, with Case Frequency, for the Purchase-to-Pay Process.

Stage 3: This stage involves enriching the *de facto* model based on the process perspective characteristics selected in stage 0, as shown in Appendix B. In addition, the *de facto* model is linked to the dashboards that are used to represent the predefined metrics using Celonis process mining software.

Stage 4: Conformance checking and process deviation analysis should be applied to combine misuse-based techniques and anomaly-based techniques. The misuse-based technique is implemented by creating dashboards that leverage business intelligence (BI) techniques for the predefined metrics, while the anomaly-based technique is implemented using the conformance checking technique.

Conformance checking is used to compare the business process with its SOP [30]. This is relevant to auditing [40] because it can detect, locate, and explain the deviation from the behavior expected in business process [56]. It helps detect the occurrence of event skipping and enables analysis of the flow of the business process [30]. Using conformance checking to classify standard and non-standard business process variants can assist in detecting potential risks [33].

A process variant is a single path (i.e., routing) that is followed by at least one business process instance [33]. All business process instances that follow the same path are grouped into the same variant [33]. Thus, it is possible to examine process variants to find out all business process instances that are in non-standard paths [33]. In turn, each process variant can be prioritized using the metrics, thereby reducing the rate of false positives in detecting fraud [25]. Reducing false positives saves time and cost [61].

Stage 5: In this stage, the fraud symptoms should be investigated with domain experts to confirm the presence or absence of fraud [25].

VI. Results and Discussion

Using the enriched model in stage 3, the conformance checking procedure was applied to extract non-conformances that form potentially fraudulent cases. The findings of the conformance checking revealed that there were 431 process flow variants (control-flow perspective). The number of variants is usually large because the process should be flexible to meet all business needs. Thus, the use of metrics as filters is essential to save time and effort, and to discover new signs of fraud.

Using the enriched model assists in fraud detection without the influence of the fraud examiner [40]. Moreover, using the predefined metrics in stage 0 ensures the accuracy and comprehensiveness of fraud detection. This is because the

[12] See www.celonis.com

predefined metrics can be used to detect fraud in the content perspective (not just the control-flow perspective) of the business process.

The combination of visual analytics and process mining can help to identify data integrity issues such as missing, non-conforming, or anomalous activities undertaken by a privileged user, or those with suspiciously short execution times [56]. Furthermore, applying the metrics using process mining reduces the number of false positives in fraud detection [25]. Thus, conformance checking and process deviation analysis are used to detect PBF [62].

In Appendices A and B, the implementation screens and results are provided. Each implementation screen serves as a link between the process flow view and the data view to present a complete view. The results show that 13 metrics produced results that should be investigated.

This implementation shows that the developed metrics can be used in the following ways: (1) directly, thereby conserving time and effort; (2) as a template, thereby facilitating the definition of other metrics and ensuring consistency among PBF detection stakeholders; and (3) to determine the implementation scope. Additionally, the developed metrics are process-oriented metrics that can measure throughput processing, as opposed to just measuring process input–output relations. This helps to detect and predict fraud, with its root cause, in its initial stages.

VII. CONCLUSIONS

This study sought to develop a comprehensive list of fraud detection metrics for business processes. A taxonomy of fraud detection metrics for business processes was used as a "building" theory to generate all possible metrics for detecting fraud in business processes. Compared to the 8 existing PBF detection metrics, 41 comprehensive metrics were developed, classified, and demonstrated. These metrics cover each of the PBF detection dimensions that are not entirely (e.g., presentation layer) or partially incorporated into existing PBF detection metrics. Additionally, their applications were demonstrated by using illustrative scenarios. Finally, their technical implementation was explained by providing an implementation that offers an accurate and comprehensive view for PBF examiners.

The study's contributions to the literature are twofold. First, the study offers improved DSR artifacts (i.e., the developed metrics and their implementation method), which can enhance the ability to detect PBF. Second, the study enriched the construction of the taxonomic theory [13] (i.e., by leveraging the taxonomy for a purpose beyond analysis). This is a step toward developing advanced theories such as design and action theory. The study also is relevant due to its practical contribution in improving PBF detection in the workplace. PBF stakeholders can improve their practices by using the developed PBF detection metrics to bolster their effectiveness.

The limited availability of data on fraud is one of the limitations of this study. This relates to the fact fraud is a sensitive topic in public discussion, and so it is not an issue spoken about openly. However, the data issued by standard-setting organizations such as the Committee of Sponsoring Organizations (COSO)[13] and the ACFE can mitigate this limitation to a certain degree. Nevertheless, the data from these organizations are mainly from the finance domain. In addition to these limitations, reviewing the metrics results with domain experts (i.e., the investigation step) is needed to confirm fraud cases. However, the scope here is specified to detect possible PBF.

Extending and validating the metrics in other domains (e.g., the telecommunications sector) is suggested as a possible direction for future research. In addition, case studies within organizations, which prioritize the use of the metrics in their specific context, are suggested. Linking each metric to a full list of possible deviation patterns is another worthwhile research opportunity. For example, the *wrong instance function* is a suitable metric that can be linked with deviation patterns such as looping, swapping, and inserting activities in the process model.

ACKNOWLEDGMENT

The authors would like to thank the deanship of scientific research at King Saud University for funding and supporting this research through the DSR Graduate Students' Research Support initiative.

REFERENCES

[1] Cotton, S. Johnigan, and L. Givarz, Fraud risk management guide. COSO, 2016.

[2] "Report to the nations: Global study on occupational fraud and abuse," 2020.

[3] D. Al-Jumeily, A. Hussain, A. MacDermott, G. Seeckts, and J. Lunn, "Methods and techniques to support the development of fraud detection system," in IWSSIP, 2015, pp. 224–227.

[4] J. J. Stoop, "Process mining and fraud detection - A case study on the theoretical and practical value of using process mining for the detection of fraudulent behavior in the procurement process," Twente University, 2012.

[5] F. Sinaga and R. Sarno, "Business process anomali detection using multi-level class association rule learning," IPTEK J. Proc. Ser., vol. 2, no. 1, 2016.

[6] D. S.Kerr, "The importance of the CobiT framework IT processes for effective internal control over financial reporting in organizations:An internationale surveys," Inf. Manag., vol. 50, no. 7, pp. 590–597, 2013.

[7] S. Huda, R. Sarno, and T. Ahmad, "Increasing accuracy of process-based fraud detection using a behavior model," Int. J. Softw. Eng. Its Appl., vol. 10, no. 5, pp. 175–188, May 2016.

[8] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, Fundamentals of business process management. New York, NY, USA: Springer, 2013.

[9] J. Jeston, Business process management practical guidelines to successful implementations. Taylor and Francis, 2017.

[10] "Fraud tree," ACFE. [Online]. Available: http://www.acfe.com/fraud-tree.aspx. [Accessed: 10-May-2020].

[11] B. Omair and A. Alturki, "A systematic literature review of fraud detection metrics in business processes," IEEE Access, vol. 8, no. 1, pp. 26893–26903, Feb. 2020.

[12] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: a comparative study," Decis. Support Syst. Sci., vol. 50, no. 3, pp. 602–613, Feb. 2011.

[13] B. Omair and A. Alturki, "Taxonomy of fraud detection metrics for business processes," IEEE Access, vol. 8, pp. 71364–71377, 2020.

[13] The COSO of the Treadway Commission is a joint initiative to combat corporate fraud. https://www.coso.org/

[14] K. Peffers, M. Rothenberger, T. Tuunanen, and R. Vaezi, “Design science research evaluation,” in International Conference on Design Science Research in Information Systems, 2012, pp. 398–410.

[15] B. Baesens, V. Van Vlasselaer, and W. Verbeke, Fraud analytics using descriptive, predictive, and social network techniques: A Guide to data science for fraud detection. Hoboken, NJ, USA: John Wiley and Sons, 2015.

[16] A. Abdallah, M. A. Maarof, and A. Zainal, “Fraud detection system: A survey,” J. Netw. Comput. Appl., vol. 68, pp. 90–113, Jun. 2016.

[17] V. Jyothsna, “A review of anomaly based intrusion detection systems,” Int. J. Comput. Appl., vol. 28, no. 7, pp. 975–8887, Sep. 2011.

[18] J. Akhilomen, “Data mining application for cyber credit-card fraud detection system,” in Industrial Conference on Data Mining, 2013, pp. 218–228.

[19] K. Mule and M. Kulkarni, “Credit card fraud detection using hidden Markov model (HMM),” Int. J. Innov. Technol. Adapt. Manag., vol. 1, no. 6, Aug. 2014.

[20] R. Nisbet, G. Miner, and K. Yale, “Fraud detection,” in Handbook of Statistical Analysis and Data Mining Applications, Amsterdam, Netherlands: Elsevier, 2018, pp. 289–302.

[21] T. W. Singleton and A. J. Singleton, Fraud risk assessment, vol. 160. John Wiley and Sons, 2011.

[22] R. C. Nickerson, U. Varshney, and J. Muntermann, “A method for taxonomy development and its application in information systems,” Eur. J. Inf. Syst., vol. 22, no. 3, pp. 336–359, May 2013.

[23] J. West and M. Bhattacharya, “An investigation on experimental issues in financial fraud mining,” in ICIEA, 2016, vol. 80, pp. 1796–1801.

[24] M. Werner, “Process model representation layers for financial audits,” Proc. Annu. Hawaii Int. Conf. Syst. Sci., vol. 2016-March, pp. 5338–5347, 2016.

[25] G. Baader and H. Krcmar, “Reducing false positives in fraud detection: Combining the red flag approach with process mining,” Int. J. Account. Inf. Syst., vol. 31, no. March, pp. 1–16, Dec. 2018.

[26] G. Vossen, “The Process Mining Manifesto—An interview with Wil van der Aalst,” Inf. Syst., vol. 37, no. 3, pp. 288–290, May 2012.

[27] H. A. Reijers, I. Vanderfeesten, and W. M. P. van der Aalst, “The effectiveness of workflow management systems: A longitudinal study,” Int. J. Inf. Manage., vol. 36, no. 1, pp. 126–141, Feb. 2016.

[28] S.-M.-R. Beheshti et al., Process Analytics. Cham: Springer International Publishing, 2016.

[29] W.-S. Yang and S.-Y. Hwang, “A process-mining framework for the detection of healthcare fraud and abuse,” Expert Syst. Appl., vol. 31, no. 1, pp. 56–68, Jul. 2006.

[30] S. Huda, R. Sarno, and T. Ahmad, “Fuzzy MADM approach for rating of process-based fraud,” J. ICT Res. Appl., vol. 9, no. 2, pp. 111–128, Nov. 2015.

[31] R. Sarno, R. D. Dewandono, T. Ahmad, M. F. Naufal, and F. Sinaga, “Hybrid association rule learning and process mining for fraud detection,” IAENG Int. J. Comput. Sci., vol. 42, no. 2, pp. 59–72, Apr. 2015.

[32] M. Jans, M. G. Alles, and M. A. Vasarhelyi, “A Field study on the use of process mining of event logs as an analytical procedure in auditing,” Account. Rev., vol. 89, no. 5, pp. 1751–1773, Sep. 2014.

[33] T. Chiu, “Exploring New Audit Evidence: the Application of Process Mining in Auditing,” Rutgers, The State University of New Jersey, 2018.

[34] M. zur Muehlen and R. Shapiro, “Business Process Analytics,” in Handbook on Business Process Management 2, Second., Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 243–263.

[35] T. Chiu, Y. Wang, and M. A. Vasarhelyi, “A framework of applying process mining for fraud scheme detection,” SSRN Electron. J., Jun. 2017.

[36] M. Leyer, D. Heckl, and J. Moormann, “Process Performance Measurement,” in Handbook on Business Process Management 2, Berlin, Heidelberg: Springer Berlin Heidelberg, 2015, pp. 227–241.

[37] M. Jans, M. Alles, and M. Vasarhelyi, “The case for process mining in auditing: Sources of value added and areas of application,” Int. J. Account. Inf. Syst., vol. 14, no. 1, pp. 1–20, 2013.

[38] C. dos S. Garcia et al., “Process mining techniques and applications – A systematic mapping study,” Expert Syst. Appl., vol. 133, pp. 260–295, Nov. 2019.

[39] R. Sarno, F. Sinaga, and K. R. Sungkono, “Anomaly detection in business processes using process mining and fuzzy association rule learning,” J. Big Data, vol. 7, no. 1, p. 5, Dec. 2020.

[40] P. Zerbino, D. Aloini, R. Dulmin, and V. Mininno, “Process-mining-enabled audit of information systems: Methodology and an application,” Expert Syst. Appl., vol. 110, pp. 80–92, Nov. 2018.

[41] W. van der Aalst et al., “Process mining manifesto,” in Lecture Notes in Business Information Processing, vol. 99 LNBIP, no. PART 1, Berlin, Heidelberg: Springer, Berlin, Heidelberg, 2012, pp. 169–194.

[42] R. Sarno and F. P. Sinaga, “Business process anomaly detection using ontology-based process modelling and Multi-Level Class Association Rule Learning,” in IC3INA, 2015, pp. 12–17.

[43] E. S. Pane, A. D. Wibawa, and M. H. Purnomo, “Event log-based fraud rating using interval type-2 fuzzy sets in fuzzy AHP,” in IEEE Region 10 Conference (TENCON), 2016, pp. 1965–1968.

[44] S. Huda, T. Ahmad, R. Sarno, and H. A. Santoso, “Identification of process-based fraud patterns in credit application,” in ICoICT, 2014, pp. 84–89.

[45] D. Rahmawati, R. Sarno, C. Fatichah, and D. Sunaryono, “Fraud detection on event log of bank financial credit business process using Hidden Markov Model algorithm,” in 3rd ICSITech, 2017, pp. 35–40.

[46] H. A. Hartanto, R. Sarno, and N. F. Ariyani, “Linked warning criterion on ontology-based key performance indicators,” in ISemantic, 2016, pp. 211–216.

[47] S. Gregor, “The nature of theory in information systems,” MIS Q., vol. 30, no. 3, pp. 611–642, Sep. 2006.

[48] J. Muntermann, R. Nickerson, and U. Varshney, “Towards the development of a taxonomic theory,” in 21st AMCIS, 2015, no. Gregor 2006, pp. 1–15.

[49] A. Dresch, D. P. Lacerda, and J. A. V. Antunes Jr, Design science research. Cham: Springer International Publishing, 2015.

[50] S. T. March, “Design and natural science research on information technology,” Decis. Support Syst., vol. 15, no. 4, pp. 251–266, Dec. 2003.

[51] O. M. Sangupamba, N. Prat, and I. Comyn-Wattiau, “Business intelligence and big data in the cloud: opportunities for design-science researchers,” in International Conference on Conceptual Modeling, 2014, pp. 75–84.

[52] A. Hevner, S. March, and J. Park, “Design science in information systems research,” MIS Q. Manag. Inf. Syst., vol. 28, no. 1, pp. 75–105, Mar. 2004.

[53] B. Omair and A. Alturki, “An improved method for taxonomy development in information systems,” Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 4, pp. 535–540, 2020.

[54] N. E. Fenton and S. L. Pfleeger, Software metrics: A rigorous and practical approach, vol. 2. 1997.

[55] A. Boenner, “Bayer: Process mining supports digital transformation in internal audit,” in Process Mining in Action, Cham: Springer International Publishing, 2020, pp. 159–168.

[56] G. Moggia and Z. Varga, “Connecting data and processes in audit — some considerations about the use of process mining,” European Court of Auditors, 2020.

[57] M. Jans, J. M. van der Werf, N. Lybaert, and K. Vanhoof, “A business process mining application for internal transaction fraud mitigation,” Expert Syst. Appl., vol. 38, no. 10, pp. 13351–13359, Sep. 2011.

[58] B. F. (Boudewijn) Van Dongen, “BPI Challenge 2019.” 4TU.Centre for Research Data, 2019.

[59] K. Diba, S. Remy, and L. Pufahl, “Compliance and Performance Analysis of Procurement Processes Using Process Mining,” in International Conference on Process Mining, 2019.

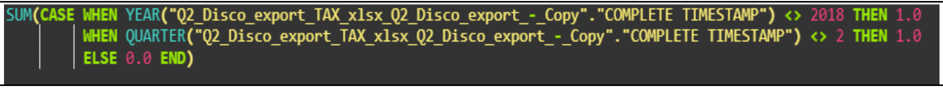
[60] W. M. P. van der Aalst, K. M. van Hee, J. M. van der Werf, and M. Verdonk, “Auditing 2.0: Using process mining to support tomorrow’s auditor,” Computer (Long. Beach. Calif)., vol. 43, no. 3, pp. 90–93, Mar. 2010.

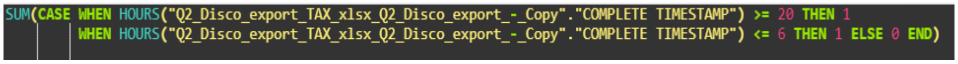
[61] J. Luell, “Employee fraud detection under real world conditions,” University of Zurich, 2010.

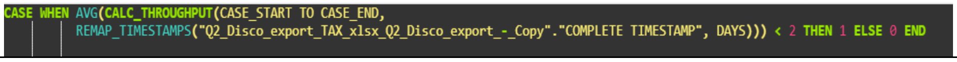
[62] G. Baader and H. Krcmar, “Reducing false positives in fraud detection: Combining the red flag approach with process mining,” Int. J. Account. Inf. Syst., vol. 31, pp. 1–16, Dec. 2018.

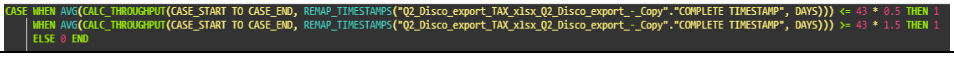
[63] L. Sánchez González, F. García Rubio, F. Ruiz González, and M. Piattini Velthuis, “Measurement in business processes: A systematic review,” Bus. Process Manag. J., vol. 16, no. 1, pp. 114–134, Feb. 2010.

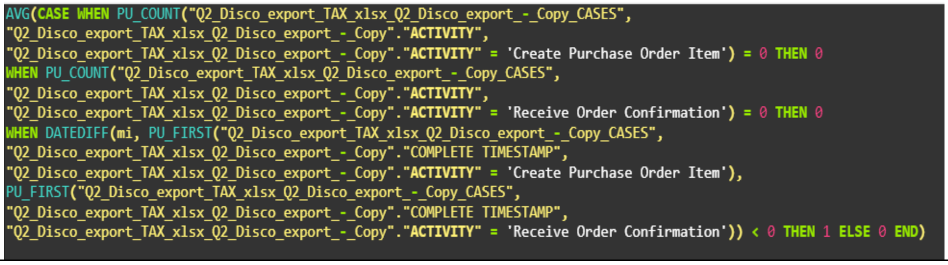
APPENDIX A

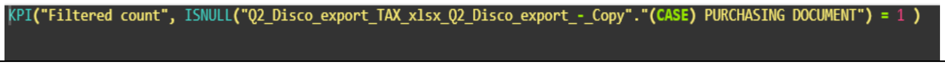
Metric name	WAT_Generic	**Threshold**	0
Fraud data scheme	Wrong	**Presentation layer**	Activity
Process perspective	Time	**Fraud domain**	Generic
Metric description	Counts the activities with execution time not in 2018 (Q2).		
Metric formula	 SUM(CASE WHEN YEAR("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP") <> 2018 THEN 1.0 WHEN QUARTER("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP") <> 2 THEN 1.0 ELSE 0.0 END)		
Result	0		

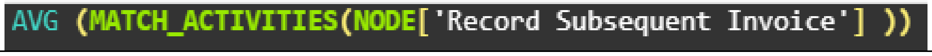
Metric name	AAT_Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Activity
Process perspective	Time	**Fraud domain**	Generic
Metric description	Counts the activities with execution time outside of normal working hours (between 8 PM and 6 AM).		
Metric formula	 SUM(CASE WHEN HOURS("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP") >= 20 THEN 1 WHEN HOURS("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP") <= 6 THEN 1 ELSE 0 END)		
Result	15254		

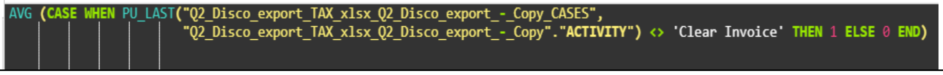
Metric name	AIT_Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Instance
Process perspective	Time	**Fraud domain**	Generic
Metric description	Monitors instances throughput time that is less than 2 days.		
Metric formula	 CASE WHEN AVG(CALC_THROUGHPUT(CASE_START TO CASE_END, REMAP_TIMESTAMPS("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", DAYS))) < 2 THEN 1 ELSE 0 END		
Result	920		

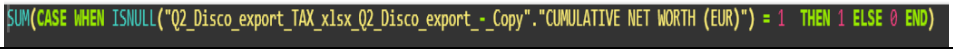
Metric name	AMT_Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Time	**Fraud domain**	Generic
Metric description	The instance throughput time is less than or greater than 43 days (average instance throughput time) by 50%		
Metric formula	 CASE WHEN AVG(CALC_THROUGHPUT(CASE_START TO CASE_END, REMAP_TIMESTAMPS("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", DAYS))) <= 43 * 0.5 THEN 1 WHEN AVG(CALC_THROUGHPUT(CASE_START TO CASE_END, REMAP_TIMESTAMPS("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", DAYS))) >= 43 * 1.5 THEN 1 ELSE 0 END		
Result	5187		

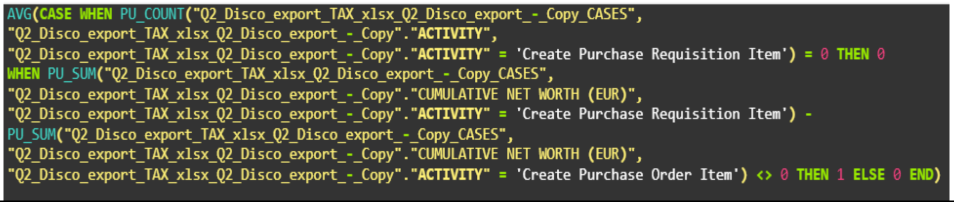
Metric name	WIF_Generic	**Threshold**	0
Fraud data scheme	Wrong	**Presentation layer**	Instance
Process perspective	Function	**Fraud domain**	Generic
Metric description	Monitors the wrong work sequence (i.e., “Create Purchase Order Item” activity occurred after “Receive Order Confirmation”)		
Metric formula	 AVG(CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Order Item') = 0 THEN 0 WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Receive Order Confirmation') = 0 THEN 0 WHEN DATEDIFF(mi, PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Order Item'), PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Receive Order Confirmation')) < 0 THEN 1 ELSE 0 END)		
Result	0		

Metric name	MIF_ Generic	**Threshold**	0
Fraud data scheme	Missing	**Presentation layer**	Instance
Process perspective	Function	**Fraud domain**	Generic
Metric description	Finds an instance where the "Purchasing Document" (PO number) is null, which is because every event should be connected to a PO number in the events log		
Metric formula	 KPI("Filtered count", ISNULL("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."(CASE) PURCHASING DOCUMENT") = 1)		
Result	0		

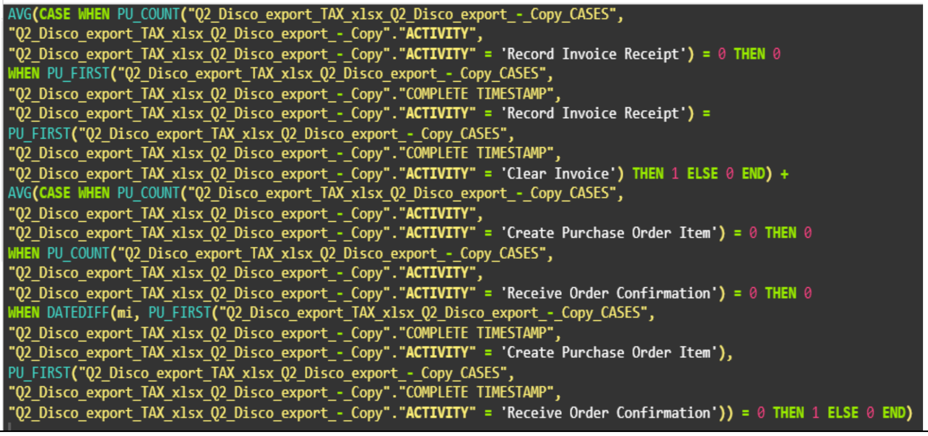
Metric name	AMF_ Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Function	**Fraud domain**	Generic
Metric description	Finds the less frequent activity "Record Subsequent Invoice" (which occurred only once in the events log)		
Metric formula	 AVG (MATCH_ACTIVITIES(NODE['Record Subsequent Invoice']))		
Result	1		

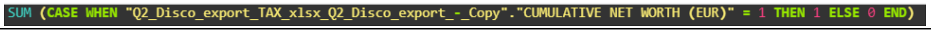
Metric name	AIF_Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Instance
Process perspective	Function	**Fraud domain**	Generic
Metric description	Counts instances that do not end with "Clear Invoice" activity		
Metric formula	 AVG (CASE WHEN PU_LAST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY") <> 'Clear Invoice' THEN 1 ELSE 0 END)		
Result	2132		

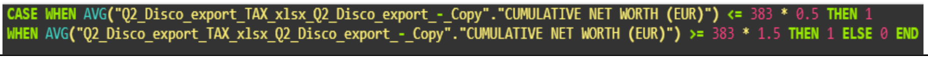
Metric name	MAD_ Generic	**Threshold**	0
Fraud data scheme	Missing	**Presentation layer**	Activity
Process perspective	Data	**Fraud domain**	Generic
Metric description	Counts activities with missing price (null).		
Metric formula	 SUM(CASE WHEN ISNULL("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)") = 1 THEN 1 ELSE 0 END)		
Result	0		

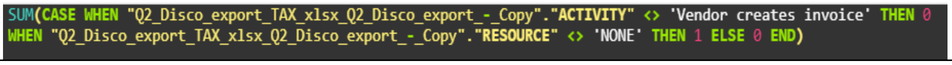
Metric name	DID_ Generic	**Threshold**	0
Fraud data scheme	Discrepant	**Presentation layer**	Instance
Process perspective	Data	**Fraud domain**	Generic
Metric description	Counts instances where the number of "Create Purchase Order Item" is not equal to that of "Create Purchase Requisition Item"		
Metric formula	 AVG(CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Requisition Item') = 0 THEN 0 WHEN PU_SUM("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Requisition Item') - PU_SUM("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Order Item') <> 0 THEN 1 ELSE 0 END)		
Result	0		

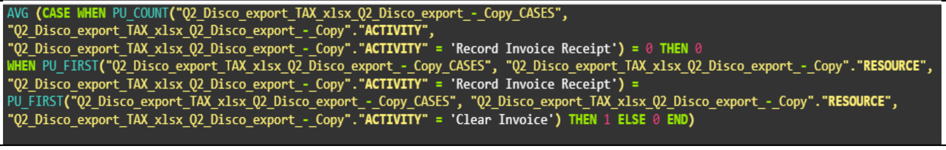
Metric name	DIF_Generic	**Threshold**	0
Fraud data scheme	Discrepant	**Presentation layer**	Instance
Process perspective	Function	**Fraud domain**	Generic
Metric description	The metric checks if the execution times of (first) "Record Invoice Receipt" and (first) "Clear Invoice" are the same, and it also checks whether "Create Purchase Order Item" activity occurred at the same time as "Receive Order Confirmation"		

Metric formula	 AVG(CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') = 0 THEN 0 WHEN PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') = PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Clear Invoice') THEN 1 ELSE 0 END) + AVG(CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Order Item') = 0 THEN 0 WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Receive Order Confirmation') = 0 THEN 0 WHEN DATEDIFF(mi, PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Create Purchase Order Item'), PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."COMPLETE TIMESTAMP", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Receive Order Confirmation')) = 0 THEN 1 ELSE 0 END)
Result	7

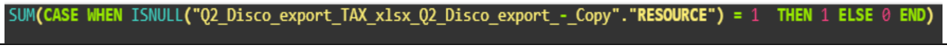
Metric name	AAD_ Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Activity
Process perspective	Data	**Fraud domain**	Generic
Metric description	Monitors activities where the price of the purchased item is equal to 1 euro		
Metric formula	 SUM (CASE WHEN "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)" = 1 THEN 1 ELSE 0 END)		
Result	13		

Metric name	AMD_Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Data	**Fraud domain**	Generic
Metric description	Counts instances that have a price less than or greater than 383 euros (average price) by 50%		
Metric formula	 CASE WHEN AVG("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)") <= 383 * 0.5 THEN 1 WHEN AVG("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)") >= 383 * 1.5 THEN 1 ELSE 0 END		
Result	10783		

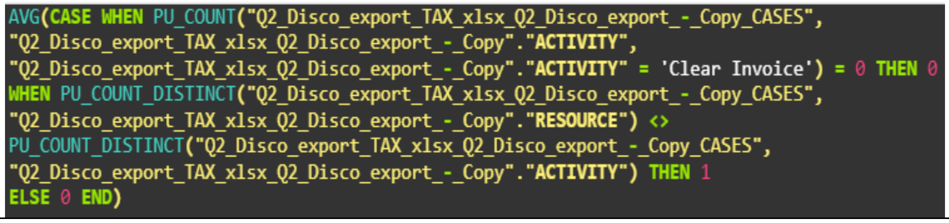
Metric name	WAR_ Generic	**Threshold**	0
Fraud data scheme	Wrong	**Presentation layer**	Activity
Process perspective	Resource	**Fraud domain**	Generic
Metric description	Monitors the resource of "Vendor Create Invoice", which should be a vendor value (i.e., NONE in the events log)		
Metric formula	 SUM(CASE WHEN "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" <> 'Vendor creates invoice' THEN 0 WHEN "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."RESOURCE" <> 'NONE' THEN 1 ELSE 0 END)		
Result	0		

Metric name	WIR_ Generic	**Threshold**	0
Fraud data scheme	Wrong	**Presentation layer**	Instance
Process perspective	Resource	**Fraud domain**	Generic
Metric description	Checks the violation of the segregation of duties rule, where the resource of "Record Invoice Receipt" should not be the same resource as "Clear Invoice"		
Metric formula	 AVG (CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') = 0 THEN 0 WHEN PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."RESOURCE", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') = PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."RESOURCE", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Clear Invoice') THEN 1 ELSE 0 END)		
Result	13		

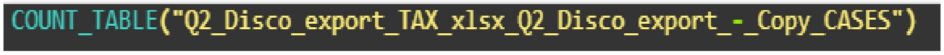
Metric name	MAR_ Generic	**Threshold**	0
Fraud data scheme	Missing	**Presentation layer**	Activity

Process perspective	Resource	**Fraud domain**	Generic
Metric description	Counts activities with null resource		
Metric formula	 `SUM(CASE WHEN ISNULL("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."RESOURCE") = 1  THEN 1 ELSE 0 END)`		
Result	0		

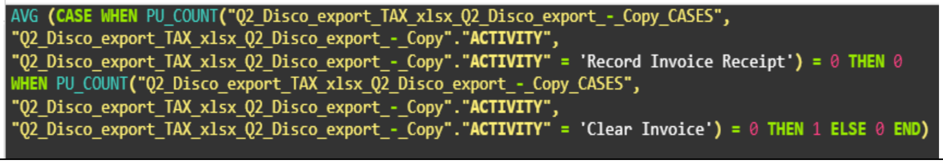
Metric name	AIR_ Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Instance
Process perspective	Resource	**Fraud domain**	Generic
Metric description	Checks whether a resource undertook more than one activity in a complete instance		

Metric formula	 `AVG(CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Clear Invoice') = 0 THEN 0`
	`WHEN PU_COUNT_DISTINCT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."RESOURCE") <>`
	`PU_COUNT_DISTINCT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY") THEN 1`
	`ELSE 0 END)`		

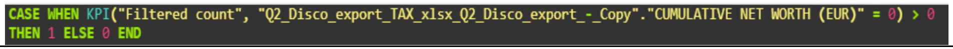
Result	963		

Metric name	AMR_Generic	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Resource	**Fraud domain**	Generic
Metric description	Monitors to determine whether employee frequency is suspicious (e.g., appears only once)		
Metric formula	 `COUNT_TABLE("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES")` (based on resource dimension)		
Result	0		

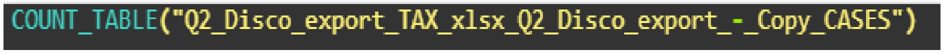
Metric name	MIF_ BillAndHold	**Threshold**	0
Fraud data scheme	Missing	**Presentation layer**	Instance
Process perspective	Function	**Fraud domain**	Finance, fictitious expenses, bill-and-hold
Metric description	By using “Missing Instance Function”, it is possible to define this specific metric, which checks whether the "Clear Invoice" activity is missing, while "Record Invoice Receipt" exists		

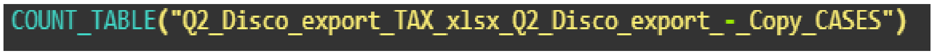
Metric formula	 `AVG (CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') = 0 THEN 0`
	`WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Clear Invoice') = 0 THEN 1 ELSE 0 END)`		

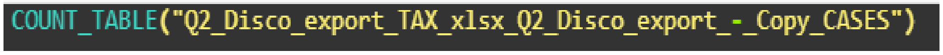
Result	4		

Metric name	WAD_ OmiOfExp	**Threshold**	0
Fraud data scheme	Wrong	**Presentation layer**	Activity
Process perspective	Data	**Fraud domain**	Finance, concealed liabilities and expenses, omission of expenses
Metric description	By using “wrong activity data”, this specific metric can be defined, which checks to see whether the activity price is equal to zero		
Metric formula	 `CASE WHEN KPI("Filtered count", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."CUMULATIVE NET WORTH (EUR)" = 0) > 0` `THEN 1 ELSE 0 END`		

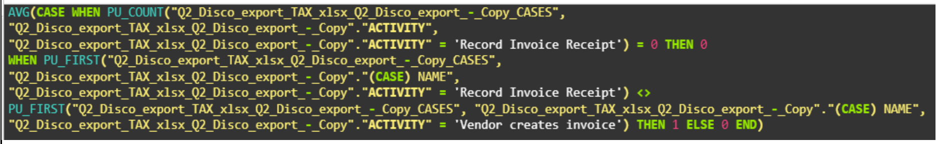
Result	237

Metric name	AMR_ ExcOfSup	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Resource	**Fraud domain**	Finance, economic extortion, exclusion of specific supplier
Metric description	Checks if the supplier frequency (e.g., #PO and PO value) is sharply decreasing over time		
Metric formula	 `COUNT_TABLE("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES")` (based on time dimension)		
Result	0		

Metric name	AMR_OwnershipOfSup	**Threshold**	N/A
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Resource	**Fraud domain**	Finance, conflict of interest, ownership of supplier
Metric description	Checks if the supplier frequency (e.g., #PO and PO value) is sharply increasing over time		
Metric formula	 `COUNT_TABLE("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES")` (based on vendor dimension)		
Result	0		

Metric name	AMR_ FictitiousSup	**Threshold**	0, 2
Fraud data scheme	Anomalous	**Presentation layer**	Model
Process perspective	Resource	**Fraud domain**	Finance, fraudulent disbursements of cash, fictitious suppliers
Metric description	By using “Anomalous model resource”, this specific metric can be defined, which checks for a supplier that appears only once (showing supplier frequency)		
Metric formula	 `COUNT_TABLE("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES")` (based on vendor dimension)		
Result	155		

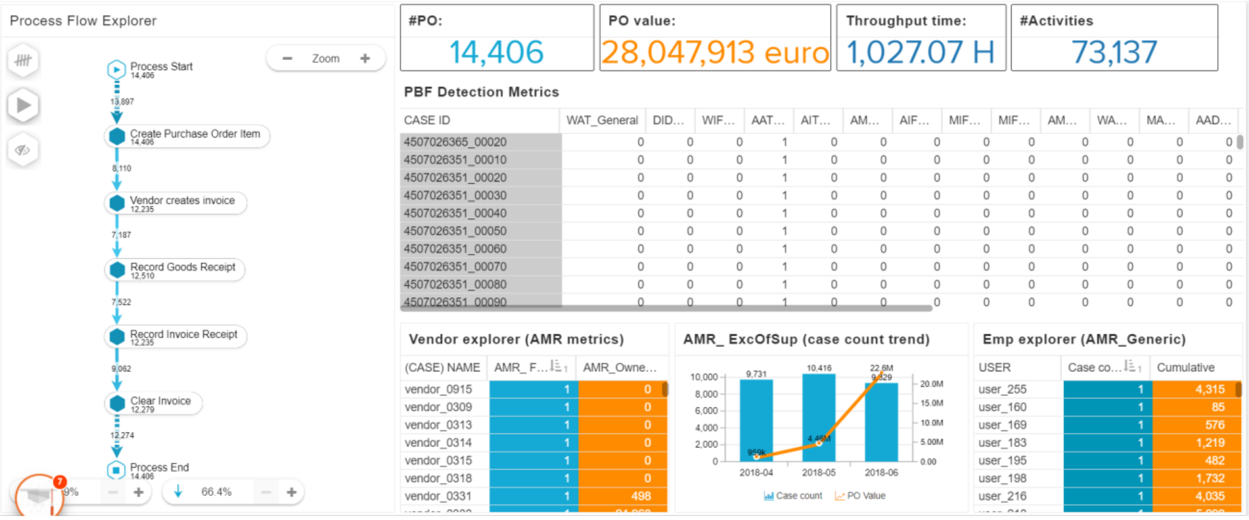
Metric name	DIR_PhantomSup	**Threshold**	0
Fraud data scheme	Discrepant	**Presentation layer**	Instance
Process perspective	Resource	**Fraud domain**	Finance, fictitious expenses, phantom supplier
Metric description	Checks if vendor in “Record Invoice Receipt” is different than that in “Vendor Create Invoice”		

Metric formula	 `AVG(CASE WHEN PU_COUNT("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') = 0 THEN 0`
	`WHEN PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."(CASE) NAME",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Record Invoice Receipt') <>`
	`PU_FIRST("Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy_CASES", "Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."(CASE) NAME",`
	`"Q2_Disco_export_TAX_xlsx_Q2_Disco_export_-_Copy"."ACTIVITY" = 'Vendor creates invoice') THEN 1 ELSE 0 END)`		

Result	0

Metric name	AIF_ FakeInv	**Threshold**	0
Fraud data scheme	Anomalous	**Presentation layer**	Instance
Process perspective	Function	**Fraud domain**	Finance, fictitious expenses, fake invoice
Metric description	By using "Anomalous instance function", this specific metric can be defined, which checks activity frequency for "Cancel Invoice Receipt" to determine whether it occurs more than once. This is because fraud may be undertaken by creating fake invoices (e.g., to increase expenses for any reason), which are canceled at a later date.		
Metric formula	`CASE WHEN KPI("Ratio", MATCH_ACTIVITIES(NODE_ANY['Cancel Invoice Receipt'] ) = 1) > 1 THEN 1 ELSE 0 END`		
Result	0		

APPENDIX B



Process Flow Explorer

#PO:	PO value:	Throughput time:	#Activities
14,406	28,047,913 euro	1,027.07 H	73,137

PBF Detection Metrics

CASE ID	WAT_General	DID…	WIF…	AAT…	AIT…	AM…	AIF…	MIF…	MIF…	AM…	WA…	MA…	AAD…
4507026365_00020	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00010	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00020	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00030	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00040	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00050	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00060	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00070	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00080	0	0	0	1	0	0	0	0	0	0	0	0	0
4507026351_00090	0	0	0	1	0	0	0	0	0	0	0	0	0

Vendor explorer (AMR metrics)

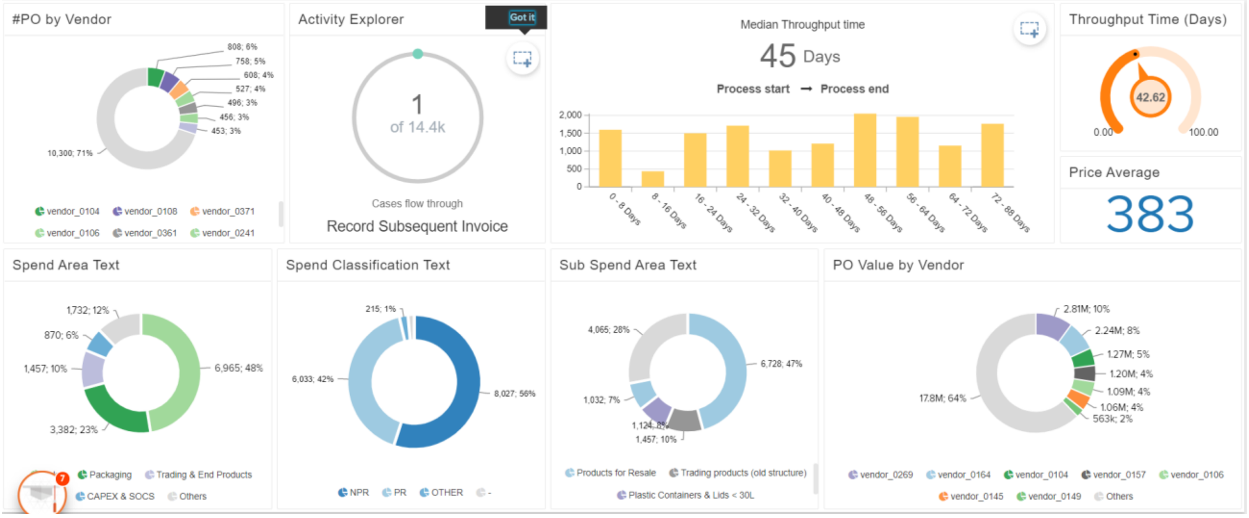
(CASE) NAME	AMR_ F…	AMR_Owne…
vendor_0915	1	0
vendor_0309	1	0
vendor_0313	1	0
vendor_0314	1	0
vendor_0315	1	0
vendor_0318	1	0
vendor_0331	1	498

AMR_ ExcOfSup (case count trend)

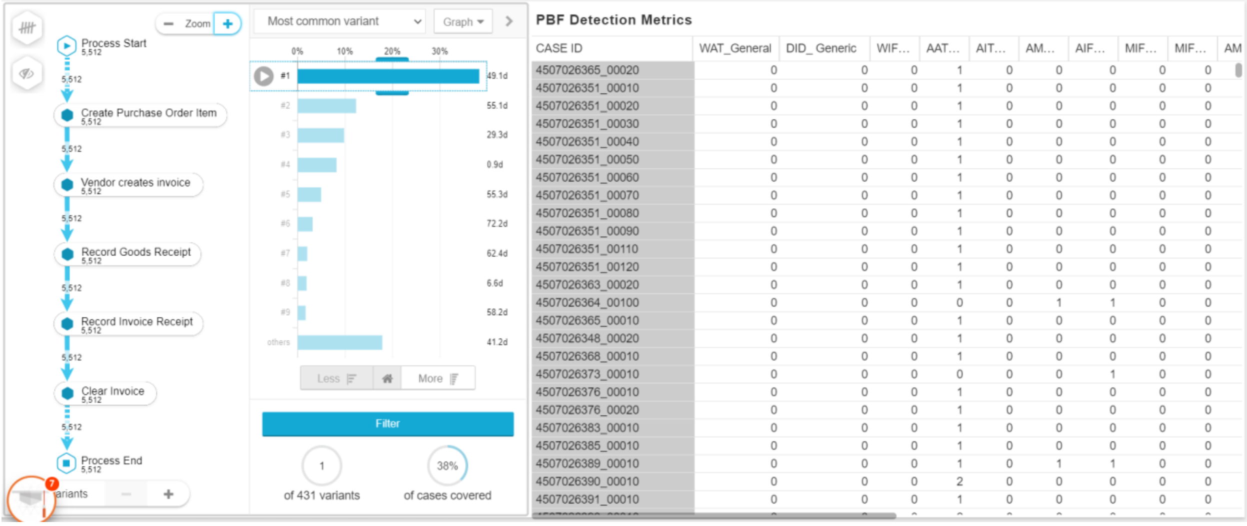
Emp explorer (AMR_Generic)

USER	Case co…	Cumulative
user_255	1	4,315
user_160	1	85
user_169	1	576
user_183	1	1,219
user_195	1	482
user_198	1	1,732
user_216	1	4,035

Screen 1. Process view Linked with Data view.



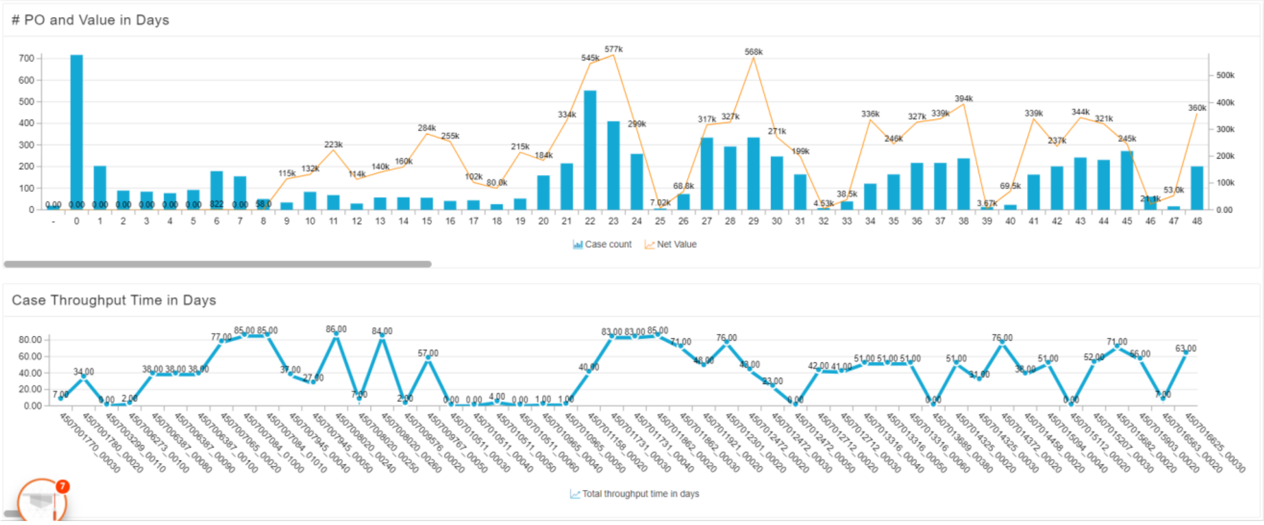
Screen 2. BI Dashboards for Analyzing Process Content.



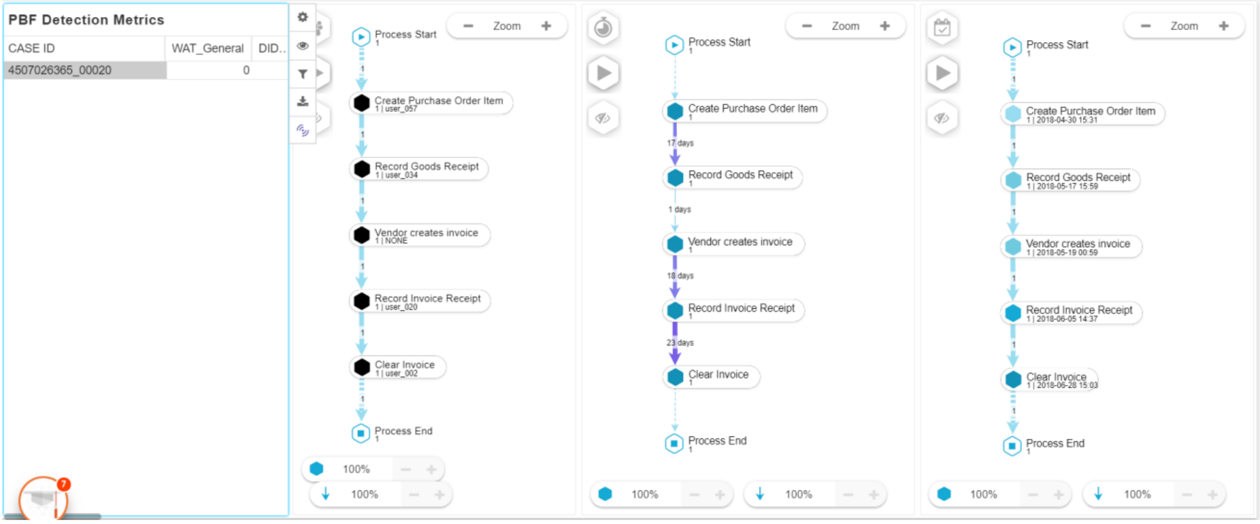
PBF Detection Metrics

CASE ID	WAT_General	DID_ Generic	WIF…	AAT…	AIT…	AM…	AIF…	MIF…	MIF…	AM
4507026365_00020	0	0	0	1	0	0	0	0	0	
4507026351_00010	0	0	0	1	0	0	0	0	0	
4507026351_00020	0	0	0	1	0	0	0	0	0	
4507026351_00030	0	0	0	1	0	0	0	0	0	
4507026351_00040	0	0	0	1	0	0	0	0	0	
4507026351_00050	0	0	0	1	0	0	0	0	0	
4507026351_00060	0	0	0	1	0	0	0	0	0	
4507026351_00070	0	0	0	1	0	0	0	0	0	
4507026351_00080	0	0	0	1	0	0	0	0	0	
4507026351_00090	0	0	0	1	0	0	0	0	0	
4507026351_00110	0	0	0	1	0	0	0	0	0	
4507026351_00120	0	0	0	1	0	0	0	0	0	
4507026363_00020	0	0	0	1	0	0	0	0	0	
4507026364_00100	0	0	0	0	0	1	1	0	0	
4507026365_00010	0	0	0	1	0	0	0	0	0	
4507026348_00020	0	0	0	1	0	0	0	0	0	
4507026368_00010	0	0	0	1	0	0	0	0	0	
4507026373_00010	0	0	0	0	0	0	1	0	0	
4507026376_00010	0	0	0	1	0	0	0	0	0	
4507026376_00020	0	0	0	1	0	0	0	0	0	
4507026383_00010	0	0	0	1	0	0	0	0	0	
4507026385_00010	0	0	0	1	0	0	0	0	0	
4507026389_00010	0	0	0	1	0	1	1	0	0	
4507026390_00010	0	0	0	2	0	0	0	0	0	
4507026391_00010	0	0	0	1	0	0	0	0	0	

Screen 3. Process Variant Explorer with Predefined Metrics.



Screen 4. Trend Analysis Dashboards.


PBF Detection Metrics

CASE ID	WAT_General	DID..
4507026365_00020	0	

Zoom
Process Start 1
1
Create Purchase Order Item 1 | user_057
1
Record Goods Receipt 1 | user_034
1
Vendor creates invoice 1 | NONE
1
Record Invoice Receipt 1 | user_020
1
Clear Invoice 1 | user_002
1
Process End 1
100%
100%
Zoom
Process Start 1
Create Purchase Order Item 1
17 days
Record Goods Receipt 1
1 days
Vendor creates invoice 1
18 days
Record Invoice Receipt 1
23 days
Clear Invoice 1
Process End 1
100%
100%
Zoom
Process Start 1
1
Create Purchase Order Item 1 | 2018-04-30 15:31
1
Record Goods Receipt 1 | 2018-05-17 15:59
1
Vendor creates invoice 1 | 2018-05-19 00:59
1
Record Invoice Receipt 1 | 2018-06-05 14:37
1
Clear Invoice 1 | 2018-06-28 15:03
1
Process End 1
100%
100%

Screen 5. An Enriched Process Model for Case id 4507026365_00020, Showing Activity Frequency, Username, Average Throughput Time, and Execution Timestamp.