IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 18, 2020, accepted April 1, 2020, date of publication April 13, 2020, date of current version April 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2987337

Taxonomy of Fraud Detection Metrics for Business Processes

BADR OMAIR AND AHMAD ALTURKI
Department of Information Systems, Faculty of Computer and Information Sciences, King Saud University, Riyadh 11543, Saudi Arabia
Corresponding author: Badr Omair (balomair@gmail.com)

This work was supported by the Deanship of scientific research through the initiative of DSR Graduate Students Research Support (GSR).

ABSTRACT A business process is a set of connected events, activities, and decision points, including actors and objects, which collectively produce a beneficial outcome for the customer. The success of an organization's strategic goals and performance depends on how well these business processes are implemented and executed. However, process-based fraud (PBF), a type of fraud that occurs in business processes, can be an obstacle to achieving this. Literature analysis shows that to date PBF detection metrics have not been sufficiently addressed. Specifically, there is overlap, confusion, and no standard for fraud definitions and categories that can affect our understanding of fraud mechanisms. This study develops a taxonomy to expose the dimensions, characteristics, and objects of PBF detection and to determine their relationships by using the design science research methodology. The developed taxonomy identifies four PBF dimensions with the following characteristics: (1) process perspective {time, function, data, resource, and location}, (2) presentation layer {process map, process stream, process model, process instance, and process activity}, (3) fraud data scheme {anomalous, discrepant, missing, and wrong}, and (4) fraud domain {generic and specific}. The objective of this taxonomy is to offer a useful tool to anyone seeking to classify, develop, and evaluate PBF detection metrics, along with a holistic view of PBF detection and the determination of its borders. Additionally, it may help in standardizing the concepts of PBF detection metrics to ensure consistency between stakeholders.

INDEX TERMS Business process fraud, fraud categories, fraud classification, fraud detection, fraud indicators, fraud metrics, fraud symptoms, fraud taxonomy, process-based fraud (PBF), red flags.

I. INTRODUCTION

Fraud can be defined as any deliberate act designed to deceive others that causes victims to suffer a loss and/or perpetrators to achieve a gain [1]. The use of one's job for personal enrichment through intentional abuse or misapplication of an employer's organization, resources, or assets is a type of fraud called occupational fraud, defined by the Association of Certified Fraud Examiners[1] [2]. *Process-based fraud (PBF)* is a type of fraud that occurs in business processes in which they deviate from the standard and normal operating procedures [3], [4]. However, in reality, not all deviations from the *standard operating procedures*[2] are fraud [3]. Expert investigations are required to confirm the occurrence of fraud.

The most noticeable outcomes of fraud in organizations are financial devastation and a tarnished reputation [5]. It is estimated that organizations typically lose 5% of their revenue owing to fraud [6]. Fraud ultimately leads to an increase in costs and also damages the customer experiences and relationships [5]. Consequently, fraud is a severe problem with far-reaching consequences [7], [8]. In 2012, there were 1,388 reported cases of fraud in 96 countries, resulting in losses of up to USD 1.4 billion [9]. However, manipulation is still ongoing, likely on a vast scale [10], and the number and volume of fraudulent incidents have been increasing [11].

Implementing fraud detection techniques can help organizations in identifying and recognizing fraud [12]. Detection can be well executed by using taxonomies because they can help in deciphering the initial list of fraudulent schemes [1], [13]. Taxonomies also contribute to the general knowledge base and research by classifying objects, thereby allowing researchers and practitioners to understand and

The associate editor coordinating the review of this manuscript and approving it for publication was Saqib Saeed.

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

[2] SOPs are a collection of documented instructions that are followed for executing a routine or repetitive activities in an organization [70]. (SOPs)

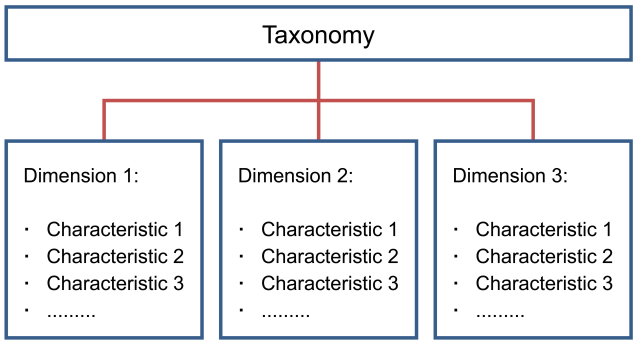

FIGURE 1. Taxonomy structure.

analyze the complex phenomena and domains, especially when little information is available about them [14], [15].

Taxonomy is a system that can be developed conceptually or empirically for grouping objects [14]. It is a set of N dimensions, each consisting of characteristics,[3] as depicted in Fig. 1. In its simplest form, taxonomy is a type of analysis theory,[4] which specifies the dimensions/characteristics of objects by describing their shared features [16]. These objects can include anything in a domain of interest that needs to be classified by taxonomy, which, in our case, are PBF detection metrics.

The literature review presented in the following section shows that there is a lack of comprehensive taxonomies of fraud detection metrics for business processes. Consequently, the following problems arise. (1) There are overlapping and frequently confusing definitions and categories of fraud that affect our understanding of its mechanisms and consequences [17]. (2) PBF terms and concepts are not standardized. (3) The relationships between PBF detection attributes are not clear. (4) The scope of PBF is not well defined. (5) The complete PBF picture is missing (which means that some PBF detection metrics are missing [18]). (6) Currently, there is no standard method of classifying the existing and new metrics because (7) taxonomy is a part of the analysis theory, which is considered the foundation (base knowledge) for obtaining knowledge [16], and improvements without taxonomy are always confusing or incomplete.

This work analyzes how detection metrics can be best classified for possible PBF into a proposed taxonomy, which can then be used to organize, simplify, and extend the PBF detection metrics. The proposed taxonomy was developed using the method presented by Nickerson *et al.* [14], which is explained in Section III.

[3]A dimension is sometimes designated as a variable, with its characteristics being the potential values (domains) of the variable [14]. However, dimensions and characteristics are common terms and can apply to all forms of taxonomies [14].

[4]Theory is an abstract entity that attempts to describe, explain, and improve the understanding of the world and, in some cases, provides predictions for the future and a basis for interventions and actions [16]. Gregor classified IS theories into five types: analytic, explaining, prediction, explaining and prediction, and design and action theory. For more information, please see [16].

The remainder of this paper is organized as follows. Section II gives a review of the fraud detection literature. Section III discusses taxonomy development and methodology. Section IV outlines the development process and implementation of the proposed taxonomy. Finally, conclusions and future work are presented in Section V.

II. LITERATURE REVIEW

A detailed systematic literature review of fraud detection metrics in business processes that includes all the relevant taxonomies is presented in [18]. It reveals that PBF detection is a topic that involves two main disciplines: fraud risk management and business process management (BPM). They are defined as follows.

1) **Fraud Risk Management** is responsible for managing all types of fraud in an organization and includes methods to prevent, identify, and respond to fraud risks [1]. This includes detecting fraud in business processes, which is used to evaluate potential fraud risks and ensure the achievement of specific business objectives [1]. Fraud risk management includes both fraud detection and fraud prevention, which are necessary to effectively combat fraud [12]. Whereas fraud detection intends to discover and recognize any fraudulent activities, fraud prevention seeks to avoid or reduce fraud. Both are independent and must be aligned and considered in fraud risk management [12].
2) **Business Process Management (BPM)** is ''a structured approach employing methods, policies, metrics, management practices, and software tools to coordinate and continuously optimize an organization's activities and processes'' [19]. It is devoted to analyzing, designing, implementing, and continuously improving the business processes of organizations [20]. A business process is a set of interconnected events, activities, and decision points that can include several actors and objects, which leads to an outcome that is of value to at least one customer [21]. Business processes include systems, data, and resources that may exist both inside and outside an organization [21]. They are performed inside a single organization and may involve cooperation with other organizations [22]. Business processes are not something conducted by organizations; rather, they form the organization's business [23]. They also determine the possible revenue and, to some extent, they form the cost profile of an organization [21] because they interact directly or indirectly with the financial accounts [24]. Owing to the importance of business processes to a business, they should be protected against any threat, including fraud [1].

A summary of the literature review is presented in Fig. 2 as a literature map. As depicted, the literature lacks an entire taxonomy of detection metrics for fraud in business processes. The literature map shows that BPM and fraud risk management are the primary domains for taxonomy, and both

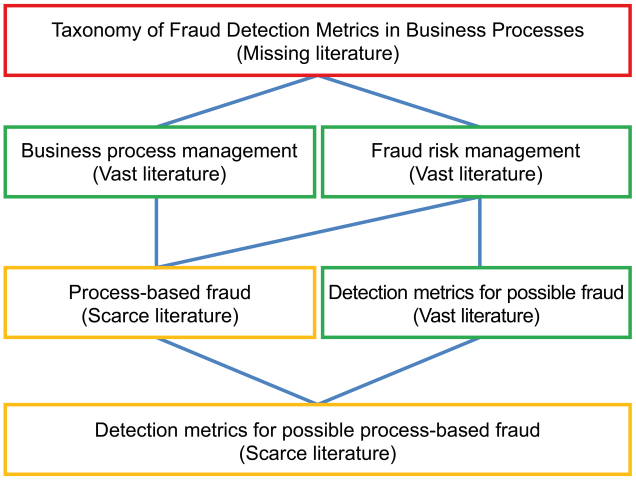

FIGURE 2. Literature map. Adapted from [18].

have received considerable attention in the literature [18]. Although a considerable amount of research has focused on the detection metrics of possible fraud, less attention has been paid to PBF and its detection metrics [18].

The literature also reveals that fraud detection techniques are generally developed based on an anomaly, misuse, or hybrid detection approaches [2]. The detection of anomalous behavior depends on detecting the deviations from the normal behavior [25]. This can help in detecting new cases of fraud, but it lacks generalization capabilities and has high rates of false alarms [26]. The misuse-based detection approach relies on the known patterns of misuse to detect any questionable transactions. It is a fast and straightforward detection technique that can be implemented by using an expert system. However, it is limited to known patterns of misuse only [27]. Finally, the hybrid approach combines anomaly-based and misuse-based fraud detection [2]. The selection of the best approach depends on the application domain and situation [28]. Notwithstanding, the anomaly-based approach is the most popular [2].

III. METHODOLOGY

Design science research (DSR) is a method that constructs and operationalizes research works performed in an academic environment or organizational context for building an artifact or recommendation [29]. It is based on a pragmatic viewpoint [30], which confirms the inability to separate utility from reality [29]. However, DSR should further contribute to the improvement of the scientific knowledge base beyond its pragmatic bias [29].

DSR has become an accepted paradigm in Information Systems (IS) research [31], [32]. According to March [33], artifacts can be categorized into one of the following categories: construct, model, method, and instantiation [29]. The creation of taxonomy is considered to be the formation of a model [14], [34].

Nickerson *et al.* [14] investigated the question of how taxonomy is constructed. They developed, presented, and

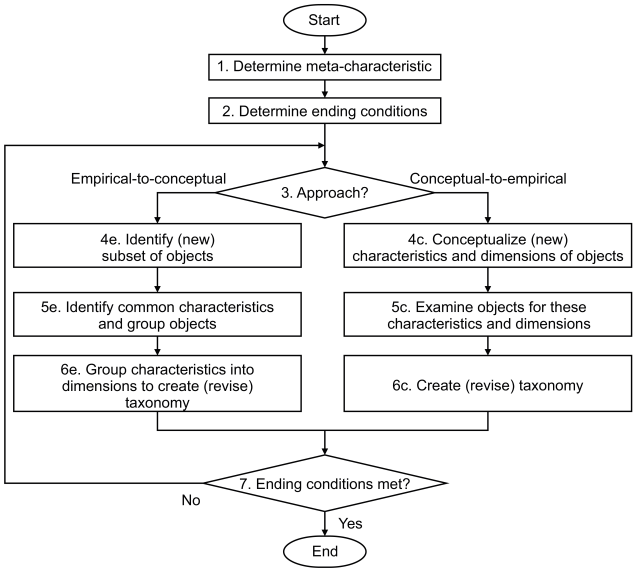

FIGURE 3. Taxonomy development method. Source: [14].

evaluated a method to develop a taxonomy that has certain qualities based on well-established literature. The method itself was built following DSR [35] by adopting and following the build/evaluate cycle for developing taxonomies and evaluating them against a set of necessary conditions [14]. The method proposed by Nickerson *et al.* [14] was employed in this study as it is suitable for developing a fraud detection metric taxonomy for business processes. This is because it involves a comprehensive systematic approach that is approved by the scientific community [14]. The method goes through several steps, as depicted in Fig. 3. They are as follows.

1) *Define the meta-characteristic of the taxonomy.* The meta-characteristic is the most general characteristic, which is the cornerstone for selecting the taxonomy characteristics. Each dimension will contain characteristics that are the logical consequences of the meta-characteristic [14]. The choice of meta-characteristic depends on the purpose of the taxonomy.
2) *Define the end conditions.* Because the method is iterative, it needs conditions to decide when to stop. The end conditions include objective and subjective conditions. The objective conditions ensure that the taxonomy satisfies its definition and it is precisely composed of dimensions, where everyone has mutually exclusive and collectively exhaustive characteristics (i.e., each dimension must have one and only one characteristic at a time). In contrast, the subjective conditions provide the researchers flexibility to add more conditions, based on their viewpoints [14].
3) *Create taxonomy using one of two approaches.* The first approach is ***conceptual-to-empirical (deductive),*** which begins with conceptualizing the dimensions of the taxonomy without considering the existing data regarding the taxonomy's objects. The conceptual

approach is based on the researcher's theory on how objects are related and how they differ. Then, the researcher uses some empirical data to determine how they match with the conceptualization to adjust the taxonomy if required. The second approach is ***empirical-to-conceptual (inductive),*** which begins with identifying the empirical data groups, followed by recognizing the nature of each group. In this approach, the researcher recognizes the characteristics of the objects that serve the meta-characteristic. Both approaches (i.e., ***conceptual-to-empirical*** and ***empirical-to-conceptual***) should be selected based on the availability of data regarding the taxonomy's objects and the knowledge of the researcher. If data access is limited and the researcher has sufficient knowledge, the *conceptual-to-empirical approach* is preferable. If data are available and the researcher has sufficient knowledge, then they may choose either approach [14].

IV. TAXONOMY

Following and implementing the steps stated in the method proposed by Nickerson *et al.* [14], as described above, the details for developing the taxonomy of detection metrics for possible fraud in business processes are as follows.

A. DEFINING META-CHARACTERISTIC (step 1 in the Nickerson et al. method)

PBF researchers, professional PBF examiners, and PBF detection technique developers are the main stakeholders in the *taxonomy of fraud detection metrics for business processes*. Therefore, improving the detection of PBF is the primary goal of all stakeholders. The taxonomy meta-characteristic is defined as *"fraud detection in business processes."*

B. DEFINING END CONDITIONS (step 2 in the Nickerson et al. method)

All objective conditions stated by Nickerson *et al.* [14] can be adopted as follows.

1) All available taxonomy objects (i.e., PBF detection metrics) must be studied.
2) No changes in the taxonomy dimensions or characteristics (i.e., adding, removing, merging, and splitting) in the last iteration can occur; if they do, then another iteration is required to examine the change impact.
3) Each characteristic for every dimension should have at least one assigned object (i.e., null characteristics should not exist).
4) All dimensions, combinations of characteristics, and all characteristics within a dimension are unique such that no duplicates exist (i.e., mutual exclusiveness).

For the subjective conditions, the following parameters are used, as suggested by Nickerson *et al.* [14].

1) **Usefulness:** the taxonomy should serve a purpose. It should have useful implications for research (rigor) and practice (relevance).
2) **Explanatory:** the taxonomy should provide valuable explanations about the nature of the existing or future objects, and their detailed attributes can be proposed. Consequently, if the characteristics of an object are known, the object can be found in a recognizable place in the taxonomy. Similarly, if an object is found in a specific place in the taxonomy, its characteristics can be identified.
3) **Conciseness:** the number of dimensions is brief yet comprehensive.
4) **Robustness:** dimensions and characteristics can be used to precisely distinguish between objects to ensure that the groups are distinct (i.e., non-overlapping groups).
5) **Comprehensiveness:** all dimensions and characteristics of the objects are identified, and there are enough parameters to classify all objects.
6) **Extendibility:** adding a new dimension or characteristic that is smooth (i.e., taxonomy is dynamic, not static).

The objective and subjective conditions that are used for evaluating the taxonomy also validate this research. This is because the validity of DSR, which is used as the methodology for this research, should emerge as a result of evaluating the developed artifact (i.e., the taxonomy) [36].

C. ITERATION 1: IMPLEMENTATION OF CONCEPTUAL-TO-EMPIRICAL APPROACH (steps 3, 4c, 5c, 6c, 7 in the Nickerson et al. method)

Because some PBF detection metrics and theoretical knowledge already exist, both approaches (i.e., conceptual-to-empirical and empirical-to-conceptual) can be used. For the first iteration, the conceptual-to-empirical approach is selected to develop the characteristics of the taxonomy. In this approach, the researcher should follow a logical process that is based on a firm theoretical foundation [37], which also includes a review of the relevant previous taxonomies [14].

Following the systematic literature review on fraud detection metrics in business processes [18], a "fraud domain" dimension that describes the area of fraud (e.g., telecommunication) can be deduced. An understanding of the fraud domain is essential for identifying the problem domain, which is a critical step in detecting fraud [38]. The fraud domain dimension has already been used to classify red flags, which are used to detect fraud [1], [39]–[41]. The fraud domain dimension can have two characteristics: "general" and "specific." The general characteristic is used to describe the metrics that can be applied to all domains, whereas the specific characteristic is used to describe the metrics that can be applied to a specific business domain (e.g., finance, insurance, telecommunication, and information technology). General and specific domain metrics are both necessary for detecting fraud in business processes; however, the taxonomy under development is not designed to focus on a

specific domain. The fraud domain can also be used to ensure that the taxonomy covers new metrics for a specific fraud area. Thus, the fraud domain as a separate dimension is added because it satisfies all objective and subjective conditions.

The second suggested dimension is the ''fraud type.'' One of the main classifications of fraud type was developed by the 2013 COSO framework [42]. It classified fraud types into categories, including fraudulent reporting, safeguarding of assets, and corruption. The fraudulent reporting category contains deliberate misstatements or omissions of amounts or disclosures to deceive people (e.g., modification of accounting records). Safeguarding of assets includes preserving the assets of the entities (e.g., property, cash) from theft, whereas corruption includes bribery and other illegal practices.

The fraud type should be determined and included in the fraud detection plan [43] because it helps to detect fraud accordingly [38], [44], [45]. However, fraud type, as a dimension, does not satisfy all the objective and subjective conditions. For example, fraud type as a dimension can have two characteristics simultaneously (e.g., fraudulent reporting and safeguarding of assets). Thus, the fraud type as a dimension is excluded.

At the end of the first iteration, the taxonomy has only one accepted dimension (i.e., fraud domain). However, a taxonomy's end conditions (i.e., both objective and subjective conditions) would still not be satisfied because this is the first iteration, and a second iteration is needed.

D. ITERATION 2: IMPLEMENTATION OF EMPIRICAL-TO-CONCEPTUAL APPROACH (steps 3, 4e, 5e, 6e, 7 from the Nickerson et al. method)

The *empirical-to-conceptual* approach can be implemented in this iteration by using PBF detection metrics available from the literature as empirical data. All distilled metrics are listed with their explanations and respective reference information in Table 1. Additionally, suggested groups were developed to cluster the metrics into common groups. After reviewing the table, a *process attribute* that refers to the process characteristics was suggested as a new dimension. The *process attribute* also has the characteristics *function*, *resource*, *decision*, and *time*, as shown in Table 1. Thus, the taxonomy now has two dimensions: *fraud domain* and *process attribute*. However, all end conditions of the taxonomy have not been satisfied yet because a new dimension is derived in this iteration. Accordingly, another iteration needs to be conducted.

E. ITERATION 3: IMPLEMENTING CONCEPTUAL-TO-EMPIRICAL APPROACH (steps 3, 4c, 5c, 6c, 7 from the Nickerson et al. method)

By examining the literature in this iteration, ''*business process perspectives*'' [51] can be considered as a replacement of the ''*process attribute*'' dimension, which was developed in iteration 2. The *business process perspectives* are more comprehensive than the *process attributes* because they include all characteristics of the *process attributes* other than ''*data*'' and ''*control-flow*.''

TABLE 1. Existing PBF detection metrics, including suggested groups. Adapted from [18].

ID	Metric name	Explanation	Reference	Suggested group
1	Skipped activity	Not performing an activity that should be executed, as stated in the SOP. The skipped activity is either a normal or decision activity [46].	[3], [4], [11], [46]–[49]	Function
2	Wrong resources	The activity is performed by an actor that is not defined by the SOP.	[3], [4], [11], [46]–[50]	Resource
3	Wrong duty	The same actor executes different activities, which should ideally have different privileges. This includes "*wrong duty sequence*" in the sequence activity, "*wrong duty decision*" in the decision activity, and "*combined wrong duty*" in the sequence and decision activities [46].	[3], [4], [11], [46]–[49]	Resource
4	Wrong pattern	Deviation from the standard sequence as stated in the SOP.	[3], [4], [11], [46]–[50]	Function
5	Wrong decision	Decision activity execution is a deviation from standard decision execution, as stated in the SOP.	[3], [4], [11], [46]–[49]	Decision
6	Wrong throughput time	Activity execution time deviates from the standard time, as stated in the SOP. It includes "*wrong throughput time min*" and "*wrong throughput time max*" [46].	[3], [4], [11], [46]–[49]	Time
7	Parallel event	Nonparallel events are performed simultaneously.	[3], [4]	Time
8	Originating behavior	The behavior of the actor while executing the activity is anomalous.	[3], [4], [48]	Resource

First, the control-flow perspective is concerned with the order of activities in the business process. Second, the resource perspective determines the action makers in process-like roles, organizational units, and authorizations. Third, the data perspective deals with the process input, consumed, and output data. Fourth, the time perspective is concerned with all process time issues, such as the execution

TABLE 2. Representation layers of business processes. Adapted from [24].

Presentation layer	Representing	Information provided on	Related information
Process map	All processes All accounts	General process structure Relationship between business processes and the BS and P&L Materiality of business processes IS supporting business processes Aggregate posting volumes created by different business processes	Process Flow Financial Statements Materiality and IS Data
Process stream	Many processes One account or group of accounts	General process structure Relationship between business processes and a selected financial account or group of accounts Materiality of business processes Aggregate posting volumes created by different business processes	Process Flow Financial Statements Materiality Data
Process model	One process Many accounts	Detailed process structure Internal controls embedded in the process Relationship between the business process and financial accounts Users involved in the process Aggregate posting volumes created by the process	Process Flow Controls Financial Statements Organizational Aspects Data
Instance model	One instance Many accounts	Detailed process structure Internal controls embedded in the process Relationship between the business process instance and financial accounts Users involved in the process Concrete journal entries and related data entries created by the process	Process Flow Controls Financial Statements Organizational Aspects Data

TABLE 3. Existing PBF detection metrics with suggested groups for the fraud data scheme dimension.

Metric name	Suggested group
Skipped activity	Missing
Wrong resources	Wrong
Wrong duty	Wrong
Wrong pattern	Wrong
Wrong decision	Wrong
Wrong throughput time	Wrong
Parallel event	Discrepant
Strange actor behavior	Anomalous

TABLE 4. Characteristics of the fraud data scheme dimension.

Fraud data scheme	Description
Wrong	Error, fictitious, non-conformity, inaccurate, duplication, and outdated aspects.
Missing	Insufficient, hidden, skipping, and lost (e.g., gaps) aspects.
Anomalous	Exceptions that are outside the normal range (e.g., an outlier, too short, too long, and excessive). It also includes ambiguity, not related, and biased aspects.
Discrepant	Inconsistency (e.g., the differences between input and output, and between past and current, discrepant parallel events).

duration and deadlines. Fifth, the function perspective describes the activities and applications of the process.

Successful detection of fraud in business processes indicates that all business process perspectives should help in detecting PBF [52]. For instance, (1) knowing the resource that posted and approved a transaction could help in detecting an unauthorized transaction or violation of duty segregation [53], (2) examining transaction activities over time may help in identifying skeptical activities such as those performed before or after off-hours [53], (3) examining the wrong process functions may indicate a fraudulent case, (4) and examining the missing data may lead to the detection of fraudulent activities.

All business process perspectives are characteristics of the process perspective as a dimension. However, the control-flow perspective merges with the function perspective because the order in which the activities are executed (i.e., the control-flow perspective) is a part of the implementation of the process activities (i.e., function perspective). Finally, location is added as a part of the process perspective because it is imperative for the auditing process [53], [54]. Moreover, to detect fraud, it is useful to know the execution site of the activity to identify the geographic risks [53].

Successful PBF needs to examine the entire business process and identify where fraud can originate [52]. Thus, the process perspective should cover all process components, which are events, activities, decision points, objects, actors, and outcomes [21]. First, the events are items that trigger the execution of activities, such as the arrival of equipment, which initiates an inspection activity. These are involved in the function and time characteristics of the process perspective. Second, the activities refer to the steps that are required to fulfill a specific work function. Third, decision refers to a particular decision made at a specific time that affects what happens later in the process, such as the approval decision. The activities and decision points are covered by the function characteristic. Fourth, the actors play roles in the process, which includes human actors, organizations, and systems. These are classified as resource characteristics. Fifth, the objects include physical objects, such as equipment, materials, and papers, as well as immaterial objects, such

TABLE 5. Evaluation of end conditions.

Condition Type	Condition Description	Explanation
Objective conditions	All available taxonomy objects must be studied.	This condition is met by reviewing all literature that mentions the PBF detection metrics, as listed in Table I. This includes reviewing the current standards for detecting fraud, such as SAS No. 99 red flags [55], as mentioned in Table VII.
	No changes (i.e., adding, removing, merging and splitting) in the taxonomy dimensions or characteristics in the last iteration.	The condition is met because the fifth iteration does not make any changes to the taxonomy.
	Every characteristic for every dimension should have at least one assigned object (i.e., null characteristic should not exist).	The condition is met because every characteristic can have at least one associated object. [5]
	All dimensions, a combination of characteristics, and all the characteristics within a dimension are unique, so no duplication exists	The condition is met because there is no duplication in all dimensions and characteristics. By using the taxonomy, full-dimensional metrics can be generated with no duplication. For example, *"wrong activity time"* in the finance domain (specific) as a PBF detection metric.
Subjective conditions	Usefulness: the taxonomy should meet its purpose. It should have useful implications for research and practice.	The usefulness condition is met when others use taxonomy [14]. However, the condition is expected to be met as the taxonomy will be used to extend the current PBF detection metrics that will help to improve the detection of PBF, as described in Section IV. For example, "*discrepant stream function*" as a metric can be generated by using the developed taxonomy. It indicates whether the execution of stream business processes contains a conflict. As an example of the possible fraud that will be exposed by this metric, they should be the same; the payment instances are more than the invoice instances as processes in the order-to-cash stream.
	Explanatory: the taxonomy should provide valuable explanations of the nature of the existing objects or future objects so that their attributes can be ascertained. Consequently, if the characteristics of an object are known, the object can be found in a recognizable place in the taxonomy. Furthermore, if an object is found in a particular place in the taxonomy, its characteristics can be identified.	The taxonomy can be used to predict the essential attributes of the PBF detection metrics, such as the implementation domain, the affected part of the business processes, and fraud data attribute. Thus, describing and classifying a new metric will be comprehensive.
	Concise: the number of dimensions is brief yet comprehensive.	The condition is met because with only four dimensions that include five characteristics at maximum, the taxonomy is still meaningful; yet, all metrics can be organized.
	Robust: dimensions and characteristics can be used to precisely distinguish between the objects so that the developed groups by the taxonomy will be distinct (i.e., non-overlapping groups).	The condition is met. Every dimension has only one characteristic at a time (i.e., mutual exclusive). Further, every metric will have unique and clear specifications.
	Comprehensive: all the dimensions and characteristics of the objects are identified. They are also sufficient to classify all objects.	The condition is met by implementing a full systematic literature review with the help of ATLAS.ti[6] software. Each dimension always has one characteristic at a time (i.e., collectively exhaustive). Finally, by using the taxonomy, the known metrics can be organized, as will be shown in VI.
	Extendible: adding a new dimension or characteristic is smooth.	The condition is met because the taxonomy is dynamic. For example, the *"fraud domain"* dimension can have specific characteristics, such as telecommunication, insurance, finance, and information technology as business domains.

as electronic records. Sixth, the outcomes are the process deliverables that are given to the customers. Both objects and outcomes are classified in the data characteristics. Therefore, the process perspective is added as a new dimension to the taxonomy and there is no need to add process components as a new dimension.

The literature reveals that business process presentation layers can also be used for auditing the process model [24].

[5]A full list of the PBF detection metrics that covers every characteristic will be presented with demonstration examples in future work owing to space limitations in this paper.

[6]https://atlasti.com

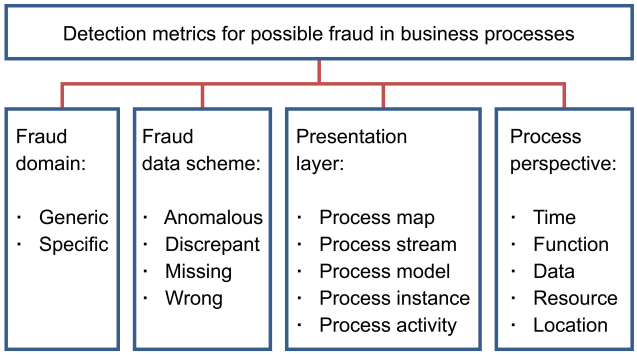

FIGURE 4. Taxonomy of fraud detection metrics in business processes.

These presentation layers include process maps, process streams, process models, and process instances [24]. The process map layer provides an overview of all processes and helps in planning and auditing the business processes. The process stream layer includes more details than the process map; it provides more information about a group of processes (e.g., procure to pay cycle) related to specific accounts. The process model layer provides detailed information/logic about the activities of an individual process whereas the process instance layer refers to a single case of the process model. This represents the actual execution of the process model, and every presentation layer is represented by specific business processes, financial accounts, and specific information. These processes are described in Table 2.

The *process activity* was also added as a characteristic of the *presentation layer* because it represents a unit of work in the process model [21]. This characteristic refers to the execution of a single activity in the process case. The *presentation layer* is now complete and is added as a new dimension to the taxonomy. Thus, the taxonomy now has three dimensions (*fraud domain, process perspective,* and *presentation layer*). However, the taxonomy's end conditions have not been satisfied yet because further new dimensions were derived during this iteration. Therefore, another iteration is still required.

F. ITERATION 4: IMPLEMENTATION OF EMPIRICAL-TO-CONCEPTUAL APPROACH (steps 3, 4e, 5e, 6e, 7 from the Nickerson et al. method)

All the applicable US Statement of Auditing Standard (SAS) No. 99 red flags [55] were used as empirical data to implement the *empirical-to-conceptual* approach in this iteration. They are the most successful and common indicators that are used to detect fraud [12]. The red flags are listed in Table 7 in Appendix A. Every red flag relevant to the detection of PBF is assigned to a developed joint group (i.e., suggested group) to identify the fraud data scheme. Moreover, all the existing metrics in Table 1 are used as additional empirical data to identify the fraud data scheme by assigning each metric to a suggested data scheme group, as shown in Table 3 . By studying the suggested groups in both tables, the *fraud data scheme* is inducted to be a new dimension; the characteristics of these dimensions are described in Table 4 .

TABLE 6. Classification of current PBF detection metrics by taxonomy.

ID	Metric name	Fraud domain	Fraud data scheme	Presentation layer	Process perspective
1	Skipped activity	Generic	Missing	Process activity	Function
2	Wrong resources	Generic	Wrong	Process activity	Resource
3	Wrong duty	Generic	Wrong	Process instance	Resource
4	Wrong pattern	Generic	Wrong	Process instance	Function
5	Wrong decision	Generic	Wrong	Process activity	Function
6	Wrong throughput time	Generic	Wrong	Process activity	Time
7	Parallel event	Generic	Discrepant	Process instance	Time
8	Originator behavior	Generic	Anomalous	Process activity	Resource

At the end of this iteration, the taxonomy has four dimensions: *fraud domain*, *process perspective*, *presentation layer*, and *fraud data scheme*. However, the taxonomy's end conditions have still not been satisfied because a new dimension has been derived during this iteration; therefore, yet another iteration is required.

G. ITERATION 5: IMPLEMENTATION OF CONCEPTUAL-TO-EMPIRICAL APPROACH (steps 3, 4c, 5c, 6c, 7 from the Nickerson et al. method)

Checking the quality of the data (i.e., if there are any missing data) is essential for detecting any performance-related issues (i.e., fraud) [56]–[58]. Thus, data quality attributes that are found in the literature can be used to determine whether the *fraud data scheme* dimension is comprehensive.

Data quality attributes were surveyed in [59], and many data quality attributes were identified. The attributes proposed in [59] are presented in Table 8 in Appendix B to check whether the characteristics of the fraud data scheme cover the relevant data quality attributes. Because all relevant data quality attributes are covered, as shown in the table, the characteristics of the fraud data scheme were not updated.

Consequently, no further changes were made to the taxonomy during this iteration, and it is expected that this iteration is the last. Thus, the end conditions for the development of taxonomy will be examined to check if they are satisfied, as described in Table 5 .

As shown in Table 5, the developed taxonomy meets all the end conditions and the taxonomy development process is complete. Fig. 4 depicts the developed taxonomy with all dimensions and characteristics.

TABLE 7. Red flags with suggested groups. Adapted from [55].

Red flag	Suggested group
Large amounts of cash on hand or processed	Anomalous
Inadequate monitoring of significant internal control	Not applicable
Inadequate segregation of duties or independent checks	Anomalous
Inadequate management oversight of employees responsible for assets	Not applicable
Inadequate access controls over automated records, including controls over and review of computer systems events logs	Not applicable
Lack of timely and appropriate documentation of transactions	Missing
Lack of complete and timely reconciliations of assets	Missing
Ineffective accounting and information systems, including situations involving reportable conditions	Not applicable
High turnover rates or employment of ineffective accounting, internal audit, or information technology staff	Anomalous
Inadequate system of authorization and approval of transactions	Not applicable
Asset, liabilities, revenues, or expenses based on significant estimates that involve subjective judgments or uncertainties that are difficult to corroborate	Anomalous
Fixed assets that are small in size, marketable, or lacking observable identification of ownership	Anomalous
Significance-related party transactions not in the ordinary course of business or with related entities that are not audited or audited by another firm	Anomalous
Inadequate physical safeguards over cash, investments, inventory, or fixed assets	Not applicable
Inadequate management understanding of information technology, which enables information technology employees to perpetuate a misappropriation	Not applicable
Significant, unusual, or highly complex transactions especially occurring close to year-end that pose difficult “substance over form” questions	Anomalous
Overly complex organizational structure involving unusual legal entities or managerial lines of authority	Anomalous
Inadequate job applicant screening of employees with access to assets	Not applicable
Inadequate record keeping concerning assets	Missing
Inventory items that are small in size, of high value, or in high demand	Anomalous
Lack of mandatory vacations for employees performing key control functions	Anomalous
Domination of management by a single person or small group in a non-owner-managed business without compensating controls	Not applicable
Ineffective board of directors or audit committee oversight over the financial reporting	Not applicable
Significant operations located or conducted across international borders in jurisdictions where differing business environments and culture exist	Anomalous
A strong financial presence or ability to dominate a certain industry sector that	Not applicable

Finally, to show the applicability of the taxonomy, Table 6 demonstrates and classifies all the current PBF detection metrics, which are listed in Table 1, into the taxonomy.

TABLE 7. *(Continued.)* **Red flags with suggested groups. Adapted from [55].**

allows the entity to dictate terms or conditions to suppliers or customers that may result in inappropriate or not arm's length transaction	
Significant bank accounts, subsidiary, branch operations in tax-haven jurisdiction for which there appears to be no clear business justification	Anomalous
Easily convertible assets, such as bearer bonds, diamonds, or computer chips	Anomalous
Difficulty in determining the organization or individuals that have controlling interest in the entity	Anomalous
High turnover of chief executive officers or board of directors	Anomalous
New accounting, statutory, or regulatory requirements	Anomalous
Significant portions of management's compensation represented by bonuses and stock options being contingent upon achieving aggressive targets for the stock price, operating results, financial position, or cash flow	Not applicable
High degree of competition or market saturation accompanied by declining margins	Not applicable
Rapid growth or unusual profitability, especially compared to that of other companies in the same industry	Anomalous
Marginal ability to meet exchange listing requirements or debt repayment	Anomalous
Need to obtain additional debt or equity financing of major research and development or capital expenditures to stay competitive	Anomalous
Significant declines in customer demand and increasing business failures in the industry or overall economy	Anomalous
High vulnerability to rapid changes in technology, product obsolescence, or interest rate	Anomalous
Perceived or real adverse effects of reporting poor financial results on significant pending transactions, such as business combinations or contract awards	Not applicable
Management and/or board of directors have personally guaranteed significant debts of the firm	Not applicable
Management and/or board of directors holding significant financial interest in the entity	Anomalous
Recurring negative cash flows from operations or an inability to generate cash flows while reporting earnings and earnings growth	Discrepant
Unrealistic profitability or trend level expectations of investment analysts, institutional investors, significant creditors, or other external parties in overly optimistic press releases or annual report messages	Anomalous
Operating losses causing threat of imminent bankruptcy, foreclosure, or hostile takeover	Anomalous
Unrealistic profitability or trend level expectation by management in overly optimistic press releases or annual report messages	Anomalous
Excessive interest by management in maintaining or increasing the entity's stock price or earnings trend	Not applicable

TABLE 8. Comparisons between data quality dimensions and characteristics of fraud data scheme dimensions.

Data quality attribute	Definition	Coverage
Timeliness	The extent to which the age of data fits the current task [63].	Covered (e.g., the wrong (time) characteristic).
Currency	The extent to which the degree of data is up-to-date [59].	Covered (e.g., the anomalous characteristic).
Consistency	The extent to which data are displayed in the same format and compatible with previous data [63].	It is covered (e.g., the discrepant characteristic).
Accuracy	The extent to which data values stored in the database correspond to real-world values [57], [64].	Covered (e.g., the wrong characteristic).
Completeness	The extent to which data are of sufficient breadth, depth, and scope for the current task [63].	Covered (e.g., the missing characteristic).
Accessibility	The extent to which data are ready for use or easily and quickly retrievable [63].	Covered (e.g., the missing characteristic).
Duplication	The extent to which duplicated data exist [65].	Covered (e.g., the wrong characteristic).
Data specification	A measure of the existence, completeness, quality, and documentation of data standards, data models, business rules, metadata, and reference data [65].	It is covered (e.g., the missing characteristic).
Presentation quality	The extent to which the data format and appearance are excellent.	Not applicable
Consistent representation	The extent to which data are presented in the standard format [66].	Not applicable
Reputation	The extent to which information is highly regarded considering the source or content [63].	Covered (e.g., the anomalous characteristic).
Safety	It is the capability of the function to achieve acceptable levels of risk of harm to people, process, property, or environment [67].	Not applicable
Appropriate amount of data	The extent to which the volume of data is appropriate for the task at hand [66].	Covered (e.g., the anomalous and missing characteristics).
Security	The extent to which access to information is restricted appropriately to maintain its security [63].	Not applicable
Believability	The extent to which information is considered true and credible [63].	Not applicable
Understandability	The extent to which data are clear without ambiguity and easily comprehended [63].	Covered (e.g., the anomalous characteristic).
Objectively (objectivity)	The extent to which information is unbiased, unprejudiced, and impartial [63].	Covered (e.g., the anomalous characteristic).
Relevancy	The extent to which information is suitable and helpful for the task at hand [63].	Covered (e.g., the anomalous characteristic).
Effectiveness	It is the capability of the function to enable users to achieve specified goals with accuracy and completeness in a specified context of use [57].	Not applicable

V. IMPLICATIONS

A. CONTRIBUTION TO KNOWLEDGE (RIGOR)

The taxonomic theory is a form of conceptual knowledge in the epistemology field of design science [60]. Gregor [16] stated that IS theories can be classified into five types: analytical, explaining, prediction, explaining and prediction, and design and action theory. According to [16], the taxonomic theory can be considered as an analysis theory (type 1), which describes or classifies specific dimensions or characteristics of individuals, groups, situations, or events by outlining the shared features found in discrete observations [16]. This answers the ''what'' question and can be used as a foundation for developing more advanced theories (e.g., explanation, prediction, explanation and prediction, and design and action) [61].

The taxonomy has a theoretical purpose for developing a taxonomic theory to solve the classification problem and increase our understanding of PBF detection. By using the Nickerson *et al.* [14] method in this study to develop the taxonomy, the taxonomic theory can be successfully defined [14].

B. CONTRIBUTION TO PRACTICE (RELEVANCE)

Taxonomy serves a practical purpose for improving and detecting PBF by using a classification system. As stated by Recker [62], the practical implications include addressing how this research changes or influences the work practices of stakeholders. The main stakeholders of this research are PBF researchers, professional PBF examiners, and the detection technique developers for PBF.

Because the taxonomic theory considers an analysis theory, which is the first step toward developing more advanced theories [16], PBF researchers can use this to develop advanced types of theories.

Professional examiners who inspect PBF in organizations can enhance the practice of PBF detection by using the

TABLE 8. *(Continued.)* **Comparisons between data quality dimensions and characteristics of fraud data scheme dimensions.**

Interpretability	The extent to which data have appropriate languages, symbols, and units and the definition is clear [66].	Covered (e.g., the anomalous characteristic).
Ease of Manipulation	The extent to which data are easy to manipulate and can be applied to different formats [66].	Not applicable
Free of error	The extent to which data are right and reliable [66].	Covered (e.g., the wrong characteristic).
Ease of use and maintainability	The extent to which data can be accessed, used, updated, maintained, and managed [65].	Not applicable
Usability	The extent to which information is clear and easily usable [68].	Covered (e.g., the anomalous characteristic).
Reliability	Capability function to maintain a specified level of performance when used in specified conditions [67].	Not applicable
Freshness	Freshness represents a family of quality factors where each one represents some freshness aspect and has its metrics [69].	Not applicable
Value-added	The extent to which information is beneficial and provides advantages from its use [63].	Not applicable
Learnability	It means the capability of the function to enable a user to learn it [67].	Not applicable
Data decay	A measure of the rate of negative change to data [65].	Not applicable
Concise	The extent to which information is compactly represented without being overwhelming (i.e., brief in presentation, yet complete and to the point) [63].	Covered (e.g., the and missing characteristics).
Consistency and synchronization	A measure of the quality of being equivalent for the data by checking how data are used in various data stores, applications, and systems, and the processes [65].	Covered (e.g., the discrepant characteristic).
Data integrity and fundamentals	The extent to which data existence, validity, structure, content, and other basic characteristics are implemented [65].	Not applicable
Navigation	The extent to which data are easily found and linked [68].	Not applicable
Usefulness	The extent to which information is applicable and helpful for the task at hand [63].	Not applicable
Efficiency	The extent to which data can quickly meet the information needs for the task at hand [63].	Not applicable
Availability	The extent to which information is physically accessible [68].	Covered (e.g., the missing characteristic).
Data coverage	The extent to which data availability and comprehensiveness are compared to the total data universe or population of interest [65].	It is covered (e.g., the missing characteristic).
Transactability	The extent to which data characteristics produce desired business transactions or outcomes [65].	Covered (e.g., the wrong characteristic).
Timeliness and availability	The extent to which data are current and available for use during a specified time frame [65].	Covered (e.g., the missing characteristic).

proposed taxonomy to develop and extend the PBF detection metrics that were probably missing in their work previously. In contrast, the developers of PBF detection techniques will be able to create more comprehensive PBF detection techniques and algorithms using a holistic taxonomy approach. Finally, the proposed taxonomy can be a useful tool for anyone interested in applying and evaluating detection metrics for PBF.

VI. CONCLUSION

This study developed a taxonomy of detection metrics for possible PBF by using DSR. The developed taxonomy includes the following dimensions with their characteristics: process perspectives {time, function, data, resource, and location}, presentation layers {process map, process stream, process model, process instance, and process activity}, fraud data schemes {anomalous, discrepant, missing, and wrong}, and fraud domains {generic and specific}.

This taxonomy serves the practical purpose of improving PBF detection in practice, while simultaneously serving the theoretical purpose of solving the classification problem and improving the understanding of PBF detection. In summary, this taxonomy can be used to (1) shed light on the main dimensions of PBF with its relationships, (2) determine the topic borders of PBF, (3) bridge knowledge gaps in PBF detection (e.g., find missing metrics), (4) standardize concepts to provide consistency between PBF stakeholders, (5) classify PBF detection metrics, (6) form a comprehensive checklist of best practices to define PBF detection metrics, and (7) pave the way for more advanced theories regarding PBF detection.

Owing to the sensitivity of fraud, the availability of fraud data is one of the limitations of this research. However, using open data can alleviate this limitation.

For future work, the developed taxonomy, which provides the necessary knowledge as a foundation, will be used to generate new PBF detection metrics. These metrics will be theoretically and empirically validated. Furthermore, case studies may be conducted where organizations extend the proposed taxonomy to a specific domain (e.g., IT security). Finally, expanding the proposed taxonomy to include prevention metrics of PBF may be conducted in future research.

APPENDIX A

See Table 7.

APPENDIX B

See Table 8.

ACKNOWLEDGMENT

The authors would like to thank the Deanship of scientific research in King Saud University for funding and supporting this research through the initiative of DSR Graduate Students Research Support (GSR).

REFERENCES

[1] D. Cotton, S. Johnigan, and L. Givarz, *Fraud Risk Management Guide*. New York, NY, USA: COSO, 2016. [Online]. Available: https://www.coso.org/Documents/COSO-Fraud-Risk-Management-Guide-Executive-Summary.pdf
[2] A. Abdallah, M. A. Maarof, and A. Zainal, "Fraud detection system: A survey," *J. Netw. Comput. Appl.*, vol. 68, pp. 90–113, Jun. 2016.
[3] 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.
[4] S. Huda, R. Sarno, and T. Ahmad, "Increasing accuracy of process-based fraud detection using a behavior model," *Int. J. Softw. Eng. Appl.*, vol. 10, no. 5, pp. 175–188, May 2016.
[5] D. Al-Jumeily, A. Hussain, A. MacDermott, G. Seeckts, and J. Lunn, "Methods and techniques to support the development of fraud detection system," in *Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Sep. 2015, pp. 224–227.
[6] Association of Certified Fraud Examiners. (Sep. 2016). *2016 ACFE Report to the Nations|Executive Summary*. [Online]. Available: https://www.acfe.com/rttn2016/docs/2016-report-to-the-nations.pdf
[7] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decis. Support Syst.*, vol. 50, no. 3, pp. 602–613, Feb. 2011.
[8] F. Sinaga and R. Sarno, "Business process anomaly detection using multi-level class association rule learning," *IPTEK J. Proc. Ser.*, vol. 2, no. 1, p. 28, Jan. 2016.
[9] Association of Certified Fraud Examiners. (2012). *Report to the Nations on Occupational Fraud and Abuse*, [Online]. Available: https://www.acfe.com/rttn-2012.aspx
[10] M. J. Nigrini, *Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations*. Hoboken, NJ, USA: Wiley, 2011.
[11] 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 *Proc. IEEE Region 10 Conf. (TENCON)*, Nov. 2016, pp. 1965–1968.
[12] 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: Wiley, 2015.
[13] T. W. Singleton and A. J. Singleton, *Fraud Risk Assessment*, vol. 160. Hoboken, NJ, USA: Wiley, 2011.
[14] 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.
[15] D. E. Strode, "A dependency taxonomy for agile software development projects," *Inf. Syst. Frontiers*, vol. 18, no. 1, pp. 23–46, Feb. 2016.
[16] Gregor, "The nature of theory in information systems," *MIS Quart.*, vol. 30, no. 3, pp. 611–642, 2006.
[17] M. Beals, M. Deliema, and M. Deevy, "Framework for a taxonomy of fraud," in *Joint collaboration of Financial Fraud Research Center at the Stanford Center on Longevity and the FINRA Investor Education Foundation*. Stanford, CA, USA: Stanford Center Longevity, 2015.
[18] B. Omair and A. Alturki, "A systematic literature review of fraud detection metrics in business processes," *IEEE Access*, vol. 8, pp. 26893–26903, 2020.
[19] A.-W. Scheer and E. Brabänder, "The process of business process management," in *Handbook on Business Process Managemen*. Berlin, Germany: Springer, 2010, pp. 239–265.
[20] J. vom Brocke and M. Rosemann, *Business Process Management*. Hoboken, NJ, USA: Wiley, 2015.
[21] M. Dumas, M. La Rosa, J. Mendling, and H. A. Reijers, *Fundamentals of Business Process Management*. New York, NY, USA: Springer, 2013.
[22] B. Wetzstein, *KPI-Related Monitoring, Analysis, and Adaptation of Business Processes*. Berlin, Germany: Stuttgart, 2016.
[23] R. Davis and E. Brabander, *ARIS Design Platform: Getting Started with BPM*. London, U.K.: Springer, 2007.
[24] M. Werner, "Process model representation layers for financial audits," in *Proc. 49th Hawaii Int. Conf. Syst. Sci. (HICSS)*, Jan. 2016, pp. 5338–5347.
[25] V. Jyothsna, "A review of anomaly based intrusion detection systems," *Int. J. Comput. Appl.*, vol. 28, no. 7, pp. 26–35, 2011.
[26] K. Mule and M. Kulkarni, "Credit card fraud detection using hidden Markov model (HMM)," *Int. J. Innov. Technol. Adapt. Manag.*, vol. 1, no. 6, p. 30, Aug. 2014.
[27] W. Wei, J. Li, L. Cao, Y. Ou, and J. Chen, "Effective detection of sophisticated online banking fraud on extremely imbalanced data," *World Wide Web*, vol. 16, no. 4, pp. 449–475, Jul. 2013.
[28] S. Mardani and H. Shahriari, "Fraud detection in process-aware information systems using process mining," in *Proc. 21st Century, E-Syst. Theno*, Dec. 2017, pp. 307–344.
[29] A. Dresch, D. P. Lacerda, and J. A. V. Antunes, Jr., *Design Science Research*. Cham, Switzerland: Springer, 2015.
[30] R. Cole, S. Purao, M. Rossi, and M. K. Sein, "Being proactive: Where action research meets design research," in *Proc. ICIS*, 2005, p. 27.
[31] S. Gregor and A. R. Hevner, "Positioning and presenting design science research for maximum impact," *MIS Quart.*, vol. 37, no. 2, pp. 337–355, Feb. 2013.
[32] W. Kuechler and V. Vaishnavi, "The emergence of design research in information systems in north america," *J. Design Res.*, vol. 7, no. 1, p. 1, 2008.
[33] S. T. March and G. F. Smith, "Design and natural science research on information technology," *Decis. Support Syst.*, vol. 15, no. 4, pp. 251–266, Dec. 1995.
[34] A. Yang and U. Varshney, "A taxonomy for mobile health implementation and evaluation," in *Proc. 22nd Amer. Conf. Inf. Syst.*, 2016, pp. 1–10.
[35] R. H. Von Alan, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quart.*, vol. 28, no. 1, pp. 75–105, 2004.
[36] Pries-Heje and Baskerville, "The design theory nexus," *MIS Quart.*, vol. 32, no. 4, pp. 731–755, 2008.
[37] N. M. Snow and J. L. Reck, "Developing a government reporting taxonomy," *J. Inf. Syst.*, vol. 30, no. 2, pp. 49–81, Jun. 2016.
[38] J. West and M. Bhattacharya, "An investigation on experimental issues in financial fraud mining," in *Proc. IEEE 11th Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2016, pp. 1796–1801.
[39] T. P. DiNapoli, *Red Flags for Fraud*. New York, NY, USA: State of New York Office of the State Comptroller, pp. 1–14, 2008.
[40] S. Y. Huang, C.-C. Lin, A.-A. Chiu, and D. C. Yen, "Fraud detection using fraud triangle risk factors," *Inf. Syst. Frontiers*, vol. 19, no. 6, pp. 1343–1356, Dec. 2017.
[41] N. Sandhu, "Behavioural red flags of Fraud—A qualitative assessment," *J. Hum. Values*, vol. 22, no. 3, pp. 221–237, Sep. 2016.
[42] *Internal control over external financial reporting: A compendium of approaches and examples*, Committee of Sponsoring Organizations of the Treadway Commission, New York, NY, USA, Sep. 2012.
[43] G. L. Gray and R. S. Debreceny, "A taxonomy to guide research on the application of data mining to fraud detection in financial statement audits," *Int. J. Accounting Inf. Syst.*, vol. 15, no. 4, pp. 357–380, Dec. 2014.
[44] A. Dal Pozzolo, O. Caelen, Y.-A. Le Borgne, S. Waterschoot, and G. Bontempi, "Learned lessons in credit card fraud detection from a practitioner perspective," *Expert Syst. Appl.*, vol. 41, no. 10, pp. 4915–4928, Aug. 2014.
[45] E. Duman and M. H. Ozcelik, "Detecting credit card fraud by genetic algorithm and scatter search," *Expert Syst. Appl.*, vol. 38, no. 10, pp. 13057–13063, Sep. 2011.
[46] R. Sarno and F. P. Sinaga, "Business process anomaly detection using ontology-based process modelling and multi-level class association rule learning," in *Proc. Int. Conf. Comput., Control, Informat. Appl. (ICINA)*, Oct. 2015, pp. 12–17.
[47] S. Huda, T. Ahmad, R. Sarno, and H. A. Santoso, "Identification of process-based fraud patterns in credit application," in *Proc. 2nd Int. Conf. Inf. Commun. Technol. (ICoICT)*, May 2014, pp. 84–89.
[48] 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.

[49] 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 *Proc. 3rd Int. Conf. Sci. Inf. Technol. (ICSITech)*, Oct. 2017, pp. 35–40.
[50] H. A. Hartanto, R. Sarno, and N. F. Ariyani, "Linked warning criterion on ontology-based key performance indicators," in *Proc. Int. Seminar Appl. Technol. Inf. Commun. (ISemantic)*, Aug. 2016, pp. 211–216.
[51] W. M. P. van der Aalst, "Business process management: A comprehensive survey," *ISRN Softw. Eng.*, vol. 2013, pp. 1–37, Aug. 2013.
[52] R. Nisbet, G. Miner, and K. Yale, "Fraud detection," in *Handbook of Statistical Analysis and Data Mining Applications*. Amsterdam, The Netherlands: Elsevier, 2018, pp. 289–302.
[53] A. Misra and V. Walden, "Proactive fraud analysis," *Intern. Audit.*, vol. 73, no. 2, pp. 33–37, 2016.
[54] T. Seyffarth, S. Kühnel, and S. Sackmann, "A taxonomy of compliance processes for business process compliance," in *Proc. BPM*, vol. 297, 2017, pp. 71–87.
[55] G. D. Moyes, R. Young, and H. F. M. Din, "Malaysian internal and external auditor perceptions of the effectiveness of red flags for detecting fraud," *Int. J. Auditing Technol.*, vol. 1, no. 1, p. 91, 2013.
[56] S. W. Tee, P. L. Bowen, P. Doyle, and F. H. Rohde, "Factors influencing organizations to improve data quality in their information systems," *Accounting Finance*, vol. 47, no. 2, pp. 335–355, Jun. 2007.
[57] C. Batini, C. Cappiello, C. Francalanci, and A. Maurino, "Methodologies for data quality assessment and improvement," *ACM Comput. Surveys*, vol. 41, no. 3, pp. 1–52, Jul. 2009.
[58] W. Eckerson, "Data warehousing special report: Data quality and the bottom line," *Appl. Dev. Trends*, vol. 1, no. 1, pp. 1–9, May 2002.
[59] F. Sidi, P. H. Shariat Panahy, L. S. Affendey, M. A. Jabar, H. Ibrahim, and A. Mustapha, "Data quality: A survey of data quality dimensions," in *Proc. Int. Conf. Inf. Retr. Knowl. Manage.*, Mar. 2012, pp. 300–304.
[60] J. Iivari, "A paradigmatic analysis of information systems as a design science," *Scand. J. Inf. Syst.*, vol. 19, no. 2, pp. 39–64, 2007.
[61] J. Muntermann, R. Nickerson, and U. Varshney, "Towards the development of a taxonomic theory," in *Proc. AMCIS*, 2006, pp. 1–15.
[62] J. Recker, *Scientific Research in Information Systems*. Berlin, Germany: Springer, 2013.
[63] R. Y. Wang and D. M. Strong, "Beyond accuracy: What data quality means to data consumers," *J. Manage. Inf. Syst.*, vol. 12, no. 4, pp. 5–33, Mar. 1996.
[64] D. P. Ballou and H. L. Pazer, "Modeling data and process quality in multi-input, multi-output information systems," *Manage. Sci.*, vol. 31, no. 2, pp. 150–162, Feb. 1985.
[65] D. McGilvray, *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information*. Amsterdam, The Netherlands: Elsevier. 2008.
[66] L. L. Pipino, Y. W. Lee, and R. Y. Wang, "Data quality assessment," *Commun. ACM*, vol. 45, no. 4, pp. 211–218, Apr. 2002.
[67] M. Heravizadeh, J. Mendling, and M. Rosemann, "Dimensions of business processes quality (QoBP)," in *BPM*, vol. 2008, pp. 80–91.
[68] S.-A. Knight and J. Burn, "Developing a framework for assessing information quality on the world wide Web," *Inf. Sci., Int. J. Emerg. Transdiscipline*, vol. 8, pp. 159–172, Aug. 2005.
[69] V. Peralta, *Data Quality Evaluation in Data Integration Systems*. Princeton, NJ, USA: Citeseer, 2006.
[70] *Guidance for Preparing Standard Operating Procedures (SOPs)*, United States Environmental Protection Agency, Washington, DC, USA, Apr. 2007. [Online]. Available: https://www.epa.gov/sites/production

BADR OMAIR was born in Riyadh, Saudi Arabia, in 1982. He received the B.S. and M.S. degrees in information systems from King Saud University, in 2004 and 2008, respectively, where he is currently pursuing the Ph.D. degree in information systems, with a focus on fraud in business processes. His research interests include business process management, fraud detection, process mining, and design science research methodology in information systems.

AHMAD ALTURKI was born in Riyadh, Saudi Arabia, in 1978. He received the B.S. degree in information systems from King Saud University, in 2000, and the M.S. and Ph.D. degrees in information systems from the Queensland University of Technology, in 2009 and 2014, respectively.

Since 2014, he has been working as an Assistant Professor at the Department of Information Systems, King Saud University. He is currently a Scientific Reviewer of two journals and two conferences. He is the author of more than ten published articles. His research interests include business process management, entrepreneurship, design research in information systems, enterprise resource planning systems, and design science research methodologies in information systems.

Dr. Alturki was a recipient of the Best Dissertation Award from the Australian Council of Professors and Heads of Information Systems (ACPHIS), in 2014.