IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 6, 2020, accepted May 31, 2020, date of publication June 8, 2020, date of current version June 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3000747

Categorization and Organization of Database Forensic Investigation Processes

ARAFAT AL-DHAQM[1,2], (Member, IEEE), SHUKOR ABD RAZAK[1], (Member, IEEE), DAVID A. DAMPIER[3], (Senior Member, IEEE), KIM-KWANG RAYMOND CHOO[4], (Senior Member, IEEE), KAMRAN SIDDIQUE[5], (Member, IEEE), RICHARD ADEYEMI IKUESAN[6], ABDULHADI ALQARNI[7], AND VICTOR R. KEBANDE[8]

[1]Faculty of Engineering, School of Computing, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia
[2]Department of Computer Science, Aden Community College, Aden 891-6162, Yemen
[3]College of Engineering and Computer Sciences, Marshall University, Huntington, WV 25755, USA
[4]Department of Information Systems and Cyber Security, The University of Texas at San Antonio, San Antonio, TX 78249-0631, USA
[5]Information and Communication Technology Department, School of Electrical and Computer Engineering, Xiamen University, Sepang 43900, Malaysia
[6]Department of Cybersecurity and Networking, School of Information Technology, Community College of Qatar, Doha 9740, Qatar
[7]Computer Science and Engineering Department, Jubail University College, Jubail 31961, Saudi Arabia
[8]Computer Science and Media Technology Department, Malmö Universitet, 20506 Malmö, Sweden

Corresponding authors: Arafat Al-Dhaqm (mrarafat@utm.my) and Kamran Siddique (kamran.siddique@xmu.edu.my)

This work was supported in part by the Research Management Center, Xiamen University Malaysia under the XMUM Research Program Cycle 3 (Grant XMUMRF/2019-C3/IECE/0006) and in part by the Research Management Center, University Technology Malaysia under the Modeling Information Security Policy Field (Grant R. J130000.7113.04E96).

ABSTRACT Database forensic investigation (DBFI) is an important area of research within digital forensics. It's importance is growing as digital data becomes more extensive and commonplace. The challenges associated with DBFI are numerous, and one of the challenges is the lack of a harmonized DBFI process for investigators to follow. In this paper, therefore, we conduct a survey of existing literature with the hope of understanding the body of work already accomplished. Furthermore, we build on the existing literature to present a harmonized DBFI process using design science research methodology. This harmonized DBFI process has been developed based on three key categories (i.e. planning, preparation and pre-response, acquisition and preservation, and analysis and reconstruction). Furthermore, the DBFI has been designed to avoid confusion or ambiguity, as well as providing practitioners with a systematic method of performing DBFI with a higher degree of certainty.

INDEX TERMS Database forensics, database forensic investigation, digital forensics, investigation process model.

I. INTRODUCTION

The use of different terminologies along with different definitions to describe exactly the same thing, object or activity can cause confusion and ambiguity [1], which does not help reasoning in a court of law. A unique terminology along with an explicit definition is usually required to inform the reader on what each term in the process model meant [2]. This is particularly useful in digital forensics where the ambiguity of terms could result in litigation failure [3]. Otherwise, the reader may be in the dark about what the author is thinking and studying. Defining exactly what each terminology means is an important part of the process construction. As a result, semantic-based conflicts that arise between two or more terminologies, must be reconciled or harmonized based on a common interpretation.

This paper discusses the redundancy and overlaps in the DBFI processes which made the DBFI field ambiguous and heterogeneous among domain investigators. Redundancy in this regard refers to the tendency of multiple components of a model (or even different models) to imply the same meaning, even within the same context (as well as within different contexts). And in a different context, such terminology presents room for ambiguous connotations which could limit the efficacy of the proposed model. The combination of these two fundamental limitations could potentially lead to

The associate editor coordinating the review of this manuscript and approving it for publication was Lo'ai A. Tawalbeh.

evidence inadmissibility in litigation [4]. Furthermore, legal adversaries tend to seek such avenues to propound grounds for evidence dismissal. Considering the volatile and dynamic nature of digital evidence, especially potential evidence in the working memory of the drive, it is essential to uniquely specify what each component entails, in a digital forensic process model. Therefore, a structured, organized, and unified investigation process in abstract categorizations is needed to address the high degree of redundancy, and ambiguity of the investigation processes among domain investigators.

A total of 40 DBFI process models were reviewed. We adapted the design science research method (DSRM) to categorize and organize the redundant and overlapping investigation processes in this reviewed literature, based on the semantic similarities in meaning or activities [5], [6]. All redundant investigation processes that have similar semantic meaning or functional meaning are organized, merged and grouped into a separate category. Hence, three categorizations, namely: i) *Planning*, *Preparation and Pre-Response Category* (PPPRC); ii) *Acquisition and Preservation Category* (APC), and iii) *Analysis and Reconstruction Category* (ARC) are proposed. It accepts the harmonizations of the tasks, activities, and terminologies of all redundant database forensic investigation processes; thereby, addressing the heterogeneity and ambiguity of the investigation processes among domain investigators.

The rest of the paper is structured as follows: Section 2 provides the background upon which the study is built and any related work. Section 3 provides the methodology. Section 4 provides discussion and analysis results. Section 5 offers a conclusion.

II. BACKGROUND AND RELATED WORKS

As discussed earlier, existing DBFI processes have varied redundant and overlapping semantics, activities, and tasks which can cause ambiguity and confusion, particularly among newer and inexperienced domain investigators.

The forensic investigation model of [7], for example, comprises the following four database investigation processes: *Suspending Database Operation*, *Collecting Data*, *Reconstructing Database*, and *Restoring Database Integrity*. A similar, albeit granular investigation methodology was proposed by Fowler *et al.* [8] which consists of seven investigation processes: *Verification*, *System Description*, *Evidence Collection*, *Timeline Creation*, *Media Analysis*, *Data Recovery*, and *String Search*. The study in [9] further proposed an investigation of a live-response model for Oracle databases which consisted of two investigation processes: *Identification Process* and *Evidence Collection*. A four-process investigation model was proposed by [10] to address MSSQL Server forensics with the following processes: *Investigation Preparation*, *Incident Verification*, *Artifact Collection*, and *Artifact Analysis*. Son *et al.* [11] presented a model to detect and investigate malicious activities in a database server, which comprises three investigation processes, namely: *Server Detection*, *Data Collection*, and *Investigation of Data Collected*. Extending the detection process model, a four-process investigation model was proposed in [12]. The processes in this model included *Collection and Preservation*, *Analysis of Anti-forensic Attacks*, *Analysis of Database Attack*, and *Preserving Evidence Report*. Additionally, *Preliminary Analysis*, *Execution*, and *Analysis* were proposed in [13].

The forensic tamper detection model of Basu [14] was designed to handle sensitive data in a MSSQL server. Specifically, in this model, the authentication and authorization mechanisms (via a SQL server) protect against external sources, but not against malicious insiders (as noted by Pavlou [15]. Similarly, a specific discovering model was proposed by [16] to reveal data theft in a database, especially when the auditing features in a database are disabled or absent. The model shows how an incident responder or a database administrator may determine that a breach has occurred in an Oracle database server in the event that there is no audit trail. The proposed model provides the *Discovering* process that consists of several concepts and activities that have a similar meaning and functioning to the proposed *Identification* process. Reference [17] presented a model that consists of several digital forensic processes. One of these processes is *Detection*, which is designed to identify covert database systems (e.g. such as those in an organization that is abused to hide evidence of illegal activities or wrong-doings within an organization).

The model of Fasan and Olivier [18] includes a reconstruction process, designed to help forensic investigators determine the presence of data of interest in the target database, including in the event that involves database modification activities that may have removed the data. Several concepts and activities in this model are similar to concepts and activities of the proposed *Artifact Analysis* process. Reference [19] proposed a forensic model to transform the data model of the database management system (DBMS) to a state appropriate for forensic investigation. This specific model provided an *Identification* process along with associated forensic methods. The data model can be viewed as the highest level of metadata which governs the way other metadata and data in the DBMS are presented to the user. A model to detect database tampering was proposed by [20]. The model includes a mechanism to provide audit logs for transaction processing systems, which can facilitate the detection of tampering in database systems. SQL coding was used to provide authentication codes on the collected data. Thus, this model implicitly provides an identification process. The *Identification* process in this model identifies audit logs, SQL triggers, hashing algorithms, extracts data from log files and ensures the collected data's validity.

In [21], the authors presented the *Triggers*, *Logfile backups*, and *Replications* techniques, in order to collect digital evidence from database systems. *Triggers* are designed to detect data modifications, *Logfile backups* are used as a regular method to collect and maintain digital evidence of database activity, and *Replication* allows the copying and distributing of data and database objects from one environment

to another. Data between databases are also synchronized for consistency. This model also includes a *Collection* process, which uses evidence collected from the three previously discussed techniques. On a similar note, a collection process model was proposed by [22]. This model was designed to facilitate the location of key evidence and achieving both evidential integrity and reliability. This model segments a DBMS into four abstract data model, data dictionary, application schema, and application data layers, which serve to separate the various levels of DBMS metadata and data.

An *Identification* process has been proposed by [23] to protect potential evidence against attackers, even when an object has been dropped and eliminated from the database system. Several resources (fixed views and tables), can be linked together to build an accurate picture of what actions the attacker took. The investigators link these fixed views and tables through SQL queries and detect attacker actions. An investigation model to assist investigators to access the data stored in MySQL, whether the user is simply unavailable or perhaps under investigation was proposed by [24]. Five common forensic investigation processes have been proposed by [25] namely, the (i) *identification*, (ii) *collection*, (iii) *preservation*, (iv) *analysis* and (v) *presentation* processes. The investigator may need to use the higher permissions of the system administrator to bypass the user's password. This model explains two methods that help achieve this goal: *Copy* and *Plainview* methods. The *Copy* method copies the system files to a new instance of MySQL and totally defeats the password protection. The *Plainview* method looks at the system files through the command line and provides a narrow view that reveals some of the data. This investigation model proposed an *Identification* process. Recently, four investigation processes have been proposed by [26] *identification*, *artifact collection*, *artifact analysis*, and *presentation and documentation*. Bria *et al.* [27] proposed five investigation processes, which are database identification, investigation, artefacts collection, analysis, and documentation. These processes have been widely explored in other domains of digital forensics, and have been proven to indeed demonstrate some tendency of overlap [2], [3]. Therefore, clearly, all DBFI processes, activities, and tasks, are required to be specific, redundant and non-overlapping. This is essentially important to solve specific investigation challenges.

III. METHODOLOGY

In this section, we will explain the criteria used in the categorization. Firstly based on the concepts introduced in [28], the categorization integrates concepts and relationships in DBFI. This allows us to undertake model collection, classification, extraction of useful concepts, concept identification, relationship identification and finally model validation. We focus on extracting processes that can help in categorization, based on semantics, ambiguity and heterogeneity among domain investigators. The adapted process comprises the following three phases:

i. Identify and Select DBFI Models
ii. Recognize and Extract Database Forensic Investigation Processes
iii. Categorize the Extracted Database Forensic Investigation Process

Phase I: Identify and Select DBFI Models

In this step, the DBFI models were identified and selected. Several DBFI models were discussed and analyzed in the literature review. Model selection for this study was based on coverage factors that were identified in previous research [28]. Wide coverage of DBFI processes that are broadly applicable is required to fulfill the aim of categorizing the investigation process. Using a coverage metric quickly provides an indication of sourced model applicability. The model is said to have a high coverage value if the model has at least two investigation processes. The model has a reduced amount of coverage value if the model only describes one DBFI process. The output of this step is twenty-two (22) common models for categorization purposes as shown in Table 1. where ID represents the first selected model (and this is the Model ID in Table 2).

Phase II: Recognize and Extract Database Forensic Investigation Processes

In this step investigation processes from the 22 models were extracted based on criteria adapted from [41], [42]:

i. Titles, abstracts, related works, and conclusions were excluded: the investigation process was either extracted from the diagram or from the main textual model.
ii. The investigation process must have a definition, activity or task; to recognize the purpose and meaning of the process.
iii. Irrelevant investigation processes not related to conducting DBFI were excluded.
iv. Include explicit and implicit investigation processes from models. As shown in Table 2 it was discovered there are seventy-eight(78) investigation processes from the 22 DBFI models. Most of these 78 investigation processes are redundant and need to be merged and grouped into a specific categorization. The next section discusses this merging process.

Phase III: Categorization of the Extracted Database Forensic Investigation Process

This phase describes how the 78 investigation processes are grouped into several categorizations based on their similarities in meaning and activities. The same approaches have been suggested by [5], [6].

The first categorization examined investigation processes from an incident response and preparation perspective. For example, the Suspension of Database Operation process in the model of [7] cuts off access to the database server for users in order to enable the capture of database activities, while the Verification and System Description processes in the model of Fowler *et al.* [8] verifies and checks database incidents, isolates the database server, confirms the incident, and documents system information such as system name,

TABLE 1. Identified and selected DBFI models.

ID	Year	Selected Models
1.	2004	System and method for investigating a data operation performed on a database [7]
2.	2005	Forensic Analysis of a SQL Server 2005 Database Server [8]
3.	2007	Oracle Forensics Live Response [9]
4.	2008	SQL Server Forensic Analysis Methodology [10]
5.	2009	Database forensic investigation based on table relationship analysis techniques [29]
6.	2009	Evidence Investigation Methodologies for Detecting Financial Fraud based on Forensic Accounting [30]
7.	2009	On metadata context in Database Forensics [31]
8.	2011	The Method of Database Server Detection and Investigation in the Enterprise Environment [11]
9.	2012	Digital Evidence for Database Tamper Detection [32]
10.	2012	Framework for Database Forensic Analysis [33]
11.	2012	A Workflow to Support Forensic Database Analysis [34]
12.	2012	On Dimensions of Reconstruction in database forensic [18]
13.	2013	Forensic Analysis of Databases by Combining Multiple Pieces of evidence [35]
14.	2014	Database Forensics: Investigating Compromised Database Management Systems [36]
15.	2014	Role of metadata in forensic analysis of database attacks [12]
16.	2014	Towards a forensic-aware database solution: Using a secured database replication protocol and transaction management for digital investigations [37]
17.	2015	Ideal log setting for database forensics reconstruction [38]
18.	2015	Database forensic analysis through internal structure carving [39]
19.	2016	A Methodology to Test the Richness of Forensic Evidence of Database Storage Engine: Analysis of MySQL Update Operation in InnoDB and MyISAM Storage Engines [13]
20.	2016	A generic database forensic investigation process model [40].
21.	2018	CDBFIP: Common database forensic investigation processes for internet of things [26].
22.	2018	Five Stages of Database Forensic Analysis: A Systematic Literature Review [27]

TABLE 2. Extracted investigation processes.

No	Model ID	Extracted Investigation Processes	Extracted process
1	M1	Suspend Database Operation, Collecting Data, Reconstructing Database, Restoring Database Integrity	4
2	M2	Verification, System Description, Evidence Collection, Timeline Creation, Media Analysis, Data Recovery, and String Search.	7
3	M3	Identification Process, Evidence Collection	2
4	M4	Investigation Preparation, Incident Verification, Artifact Collection, Artifact Analysis	4
5	M5	Database Connection Environment, Table Relationship Search and Join Process and Data Extraction	3
6	M6	Data Acquirement with Seizure and Search Warrant, Begging of Investigation, and Financial and Business Data Analysis	3
7	M7	Metadata Extraction, Restoration and Searchability	2
8	M8	Server Detection, Data Collection, Investigation on Data Collected	3
9	M9	Setup Evidence Collection Server, Collecting Files, and Analysis Process	3
10	M10	Identification, Artifact Collection, Artifact Analysis, Final Forensic Report	4
11	M11	Incident Reporting, Examination Preparation, Physical & Digital Examination, Documentation & Presentation, Post Examination, and Post Examination Analysis Process	6
12	M12	Determine Database Dimension, Determining Acquisition Method, Collection of Volatile Artifacts, Collection of Non-volatile Artifacts, Preservation and Authentication of Collected Data, and Analysis of Collected Data	6
13	M13	Artifact Collection, Forensic Analysis	2
14	M14	Identification Process, Collect Suspect Database System	2
15	M15	Collection and Preservation, Analysis Anti-forensic Attacks, Analysis Database Attack, and Preserving Evidence Report	4
16	M16	Collection Process, Analysis Process	2
17	M17	Analysis Process and Reconstruction Process	2
18	M18	Reconstructing Volatile Artifacts, Recovering Database Schema	2
19	M19	Preliminary Analysis, Execution, Analysis	3
20	M20	Identification, Collection, Preservation, Analysis, Presentation	5
21	M21	Identification, Artifact Collection, Artifact Analysis, Documentation and Presentation	4
22	M22	Database Identification, Investigation, Artifacts Collection, Analysis, Documentation	5
Total			78

serial number, operating system, system function, and physical description. In addition, the Identification process in [43] model provides for disconnecting the database server from the network in order to capture volatile data. Similarly, the Investigation Preparation and Incident Verification processes in [10] model are used to identify and verify database incidents, begin a preliminary investigation, prepare workstations and tools for incident response, and disconnect the database server.

Furthermore, the Database Connection Environment process in the model proposed by [29] prepares the investigation environment and obtains the necessary permissions to be able to access the database and execute the required commands. Also, the purpose of the Table Relationship Search and Join process is to extract table-spaces in the database, select the target, select the tables which store investigation data, and repeatedly check the other table field.

The Data Acquirement with Seizure and Search Warrant process requires securing the incident scene and extracting evidence that relates to a crime or an incident [30]. Another process is the Server Detection process used to detect any server hosting a database system. This process includes understanding the overall network inside a company; and acquiring the network's topology to identify and detect the victim database server [11]. The Setup Evidence Collection Server process described in the [32] model is used to prepare the environment to store recorded incidents, while the Identification process described in [33] identifies relevant database files (text files, log files, binary files) and utilities.

Similarly, [34] proposed an Incident Reporting and Examination Preparation processes, which are used to capture database incidents through user reports, system audits, and/or triggered events. Database incidents are then handled by cutting off the network, configuring the investigation environment, identifying violated policies, preparing the proper tools and informing the decision-maker. In addition, [18] suggested Determining Database Dimension and Acquisition Method processes, which are used for identifying which dimension of the database has been attacked or hacked. Once this has been achieved, the proper acquisition methods for that dimension are then identified. Also, the Choose Environment and Select Implement Method process proposed by [36] is used to select the forensic environment (clean or discovered environment), and select a method that is used to transform the forensic setting into the selected forensic environment. Also, the Preliminary Analysis process is proposed by [13] that aimed to create an architectural visualization of the database with all the components and their location within the layered model of the DBMS, identify files and folders in layers below the storage engines' layer, prepare and use forensic tools and procedures to create an initial image and then collect metadata values of the identified target files, and record the metadata of the target files. The Identification process is offered by [40] that intended to prepare laws and regulations, investigation techniques, investigation team, policies, database resources, investigation environment, authorization, detection server, interview, detection database incident, and incident report. Also, the Identification process proposed by [26] is used to prepare a clean database forensic investigation environment and trusted forensic techniques, as well as allow the investigation team to isolate the database server from the network to prevent users from tampering with it, and to capture volatile and non-volatile data. Finally, [27] introduced a Database Identification process useful for defining, identifying, preparing, detecting, and investigating database incidents. This is the initial process of an investigation to find a problem in the database. This can help to identify the investigation methods to be used in this investigation process.

Thus, twenty-one (21) investigation processes have been organized and merged in the first category based on their similar activities or meaning as shown in Table 3.

The second categorization focused on data collection. For example, the *data collection* process of [7] focuses on assembling data, metadata and intruder activities. The *Evidence Collection* processes in [8], [43] are designed to collect evidence from the database server(s) of interest. The *Artifact Collection* process in [10] is designed to facilitate the collection of volatile and non-volatile MSSQL Server database artifacts such as log files, data files, data cache, transaction logs, and log files. In [29], the *Data Extraction* process allows one to extract data on relationships that are connected to columns in database tables of interest. In addition, the earlier phase of [30] investigation process has similar activities designed to acquire fraud data from the database server of relevance. Metadata Extraction process in [45] allows one to extract the metadata of the database dimension and determine the individual authorized to perform a certain action. *Data Collection* process in [11] comprises two stages, one

TABLE 3. Category A of DBFI process.

No	Similar Processes	Activity and Meaning
1.	Suspend Database Operations	Database operations are suspended, at least long enough to capture evidence of the intruder's actions. This may entail disabling new logins, terminating any or all existing sessions and disconnecting users from the database.
2.	Verification	Verifies and checks incident, isolates database server and confirms the incident.
3.	System Description	Documents system information identified in the verification processes, i.e. system name, serial number, operating system, system function, and physical description.
4.	Identification process	Disconnects database server from network to capture volatile data as well as prepare forensic environment and forensic techniques used to move captured data.
5.	Identification process	Used to disconnect the database servers from the network to capture volatile data as well as prepare forensic environment and forensic techniques used to move captured data.
6.	Investigation Preparation	Identifies and prepares forensic workstations and forensic toolkits to respond to an incident and then disconnects from the database server.
7.	Incident Verification	Verifies the database incident through preliminary investigation.
8.	Database Connection Environment	Prepares the investigation environment and obtains the right to access the database and execute the command.
9.	Table Relationship Search and Join Process	Used to extract all table-spaces in the database, select the target, select the tables which store investigation data, and repeatedly check the other table field.
10.	Data Acquirement with Seizure and Search Warrant	Detect and secure database system resources and gather evidence that relates to accounting fraud. Protects the data resources of the corporation. Also, conduct an interview with DBA to validate the existence of a server managed by the corporation.
11.	Server Detection	Server detection is used to identify and detect the victim database server.
12.	Setup Evidence Collection Server	Preparing the investigation environment to reveal an incident.
13.	Incident reporting	capture database incident through user report, system audit, or triggered events
14.	Examination Preparation	Used to detect a database incident, isolate a network, configure an investigation environment, identify policies, and prepare proper forensic tools as well as making decisions about next steps.
15.	Determine Database Dimension	Identifies which dimension of the database has been attacked or hacked
16.	Determining Acquisition Method	Identifies the proper acquisition methods for that dimension.
17.	Identification	Prepare the database forensic layers, methods and environment
18.	Preliminary analysis	Create an architectural visualization of the database, identify files and folders in layers below the storage engine layer, prepare and use forensic tools and procedures to create an initial image, collect and record metadata values of the identified target files.
19.	Identification	Used to prepare laws and regulations, investigation techniques, investigation team, policies, database resources, investigation environment
20.	Identification	Used to prepare clean database forensic investigation environment and trusted forensic techniques; Allows the investigation team to isolate the database server from the network to prevent users from tampering with and capturing volatile and non-volatile data.
21.	Database Identification	Introduced to define, identify, prepare, detect, and investigate database incidents. This is the initial process of an investigation to find out a problem in the database. This can help to find the investigation methods to be used in the investigation.

dedicated to selective files and the another for collecting entire files. The *file collection* process of [32] allows one to collect Oracle files from specific locations, prior to relocating to the evidence collection server for further investigation. The *Artifact Collection* process was also proposed in [33] to collect and extract database files and metadata from compromised MySQL Server databases. Similarly, the authors of [34] proposed a *Collection* process as a sub-process of physical and digital examination to collect physical and digital data. The *Collection of Volatile Artifacts* and *Non-Volatile Artifacts* processes were proposed in [18] to collect database files, log files, log transactions and also volatile artifacts such as data caches, redo logs, and undo logs. Similar to the *Artifact Collection* process proposed in [10], and *Artifact Collection* process proposed in [35]. The *Collect Suspect Database System* proposed in [36] allows investigators to collect and extract suspected database management system data and move it to a secure area for further forensic investigation. The *Collection and Preservation* process proposed in [12] allows investigators to collect detailed multiple logs of SQL, MySQL and operating systems. The *Collection* process that was proposed in [37] is used to gather evidence by replicating sources. The *Execution* process proposed in [13] allows investigators to use forensic tools and procedures to create forensic values and then collect metadata values of the identified target files.

Thus, twenty (21) investigation processes have been organized in a second categorization based on their similar activities or meaning as shown in Table 4.

The third categorization is broadly focused on database reconstruction, analysis, and overall forensic analysis. For example, an *Analysis* process has been proposed in several models. In the model of [7], it was used to reconstruct the database and restore database integrity after collecting data to rebuild intruder activities along with revealing malicious actions and restoring database consistency. In addition, Fowler identified it with different names in two models. For example, in the model of [8], it was mentioned as part of the *Timeline Creation, Media Analysis*, *Data Recovery* and *String Search* processes, while the authors of [10] described it as part of the *Artifact Analysis* process to reconstruct timeline events and analyze malicious activity. In addition, the model in [30] referred to the analysis process as *Financial and Business Data Analysis* and used it to reveal fraudulent transactions. Other models referred to the analysis process as *Restoration and Searchability* [31], and *Investigation on Data Collected* [11]. Furthermore, the authors of [33] mentioned it explicitly as *Artifact Analysis*, as part of the *Reconstruction* process along with the *Physical and Digital Examination* process of the model in [34]. Also, the *Forensic Analysis* process proposed by the authors of [38] uses log analysis and/or log management tools to enhance the analysis of the volume of information that may be retrieved from log files during database forensics. Other models referred to the *Analysis* process as *Forensic Analysis* [35], *Analysis Anti-Forensic Attacks*, and *Analysis Database Attack* [12], *Reconstructing Evidence* [37], *Reconstruction* [38], and *Reconstructing Volatile Artifacts* [39].

TABLE 4. Category B of DBFI process.

No	Similar Processes	Activity and Meaning
1.	Collecting Data	Assembles data, metadata and intruder activities from the database server.
2.	Evidence Collection	Collects evidence from the victim database server.
3.	Collection process	Uses collected volatile data from a compromised database server.
4.	Artifact Collection	Used to collect volatile and nonvolatile MSSQL server database artifacts such as data files, data cache, transaction logs, and log files.
5.	Data Extraction Process	Used to extract data from database tables that are identified in the Table Relationship Search and Join Process. Also collects various file types, i.e. email attachments, multimedia and image files stored in the file server, and in the database.
6.	Begging of Investigation	Extracts fraud data from the database server.
7.	Metadata Extraction	Used to extract the metadata of database dimensions used to determine who was authorized to perform a certain action.
8.	Data Collection	Divided into a stage of selectively collecting files and a stage of collecting the entire files.
9.	Collecting Files	Used to gather data from the specified sites such as redo logs, data blocks, audit trails, live response, views, oracle recycle bin, and system change number.
10.	Artifact Collection	Used to collect and extract database files and metadata from compromised databases.
11.	Collection process	Collects physical and digital data.
12.	Collection of Non-Volatile Artifacts	Collects nonvolatile artifacts such as database files, log files, and log transactions.
13.	Collection of Volatile Artifacts	Collect volatile artifacts such as data caches, redo log and undo log
14.	Artifact Collection	Used to collect volatile and nonvolatile MSSQL server database artifacts such as log files, data files, data cache, transaction logs, log files and so on
15.	Collection suspect database system	Used to collect and extract suspected database management system data and move it for further examination.
16.	Collection and preservation	Used to collect detailed logs of SQL, MySQL, and the operating system.
17.	Collection process	Gathers evidence by replicating sources
18.	Execution	Used to create a forensic image and then collect metadata values of the identified target files.
19.	Collection, Preservation	Collects and preserves evidence from database server.
20.	Artifact Collection	Collects volatile and nonvolatile artifacts
21.	Artifacts Collection	Collects volatile and nonvolatile artifacts

Therefore, twenty-one (21) investigation processes have been organized and merged based on their activities and meaning. Table 5 presents the third categorization of organized and merged investigation processes with similar meanings and activities.

IV. RESULTS AND DISCUSSION

Seventy-eight (78) common investigation processes have been extracted from 22 DBFI models. Clearly, the extracted processes were overlapping and redundant. Thus, the categorization of these processes was applied, to solve the

TABLE 5. Category C of DBFI process.

NO	Similar Processes	Activity or Meaning
1.	Reconstructing Database	Used to rebuild intruder activities and reveal malicious actions.
2.	Restoring Database Integrity	Used to restore database consistency.
3.	Media Analysis	Focuses on analyzing activities and revealing malicious intruders.
4.	Timeline Creation	Used to construct an initial timeline that maps out the notable digital events which will be used during the *Media Analysis* process
5.	Data Recovery	Recovers data to be ready for user access
6.	Search String	Used to further investigation into transactions that occurred outside of the scope of this investigation to identify rows for reconstruction.
7.	Artifact Analysis	Focuses on analyzing authentication and authorization artifacts as well as configuring and versioning artifacts. Furthermore, it analyzes activity reconstruction and data recovery artifacts, which makes up the largest grouping of artifacts.
8.	Financial and Business Data Analysis	Deep analysis of the account data and other related business data should be executed. It is used to reveal fraudulent transactions.
9.	Restoration and Searchability	Recreation of data that has been (partially) destroyed or only partially recovered.
10.	Investigation on Data Collected	Methods of investigating data can be largely divided into three types, such as investigating data collected by using an agent remotely, investigating data collected by using backup commands and investigating data collected by using the entire files.
11.	Artifact Analysis	All data acquired through incident verification and collection phases are consolidated and analyzed.
12.	Reconstruction	Rebuilds database events.
13.	Reconstruction of the database	;Rebuild database events without using log files or system metadata.
14	Forensic Analysis	Involves temporal detection, the determination of the time. It also involves spatial detection, the determination of where the location of the data in the database was altered.
15	Analysis anti-forensic attacks, Analysis database attack	Reconstruct and analyze database attacks to reveal the attacker, when the attack happened, where it happened and how it happened.
16	Reconstructing Evidence	Used to rebuild user activities and detect malicious activities.
17	Reconstruction process	Identifies changes made to a database, identifies who may be responsible for the changes, confirming what is expected to be in the database, and determines the timeline of events in the database.
18	Analysis Process	Use of log analysis and/or log management tools to enhance the analysis of the volume of information that may be retrieved from log files during database forensics; the analysis process should be automated.
19	Reconstructing Volatile Artifacts	Recover the newly introduced data from inserts and updates. Also, recover recently performed user actions (i.e., reconstructing the fact that data were inserted, deleted or updated). Additionally, discover information about changes that were canceled and undone (i.e., aborted transactions).
20	Recovering Database Schema	Discovering the original schema, structure identifiers identifies pages from the same structure, and discover other components of the schema.
21	Analysis stage	Compare the values of metadata of each file before the database operation and after the operation, identify changes in metadata values after the operation, and identify the files that have been affected.

heterogeneity and ambiguity of these overlapping and redundant processes. Thus, the categorization procedure did not rely solely on naming conventions but relied on similarities in the activities or meaning. Thus, three main categorizations have been proposed in this study which are PPPRC, APC, and ARC. Each category includes similar activities, tasks, meanings, and purposings regardless of the naming of investigation processes.

A. CATEGORY A: PLANNING, PREPARATION, AND PRE-INCIDENT RESPONSE CATEGORY (PPPRC)

This category contains 21 investigation processes that may be used for planning, preparation and database pre-incident responding. PPPRC is incorporated as a proactive approach before incident identification – see ISO/IEC 27043 [48]. According to ISO/IEC 27043, forensic readiness is optional and hence, our study does not include readiness as a mandatory process. However, in PPPRC, both preparation and planning can be used when the need arises. The entire PPPRC category's processes are used to prepare a clean database forensic investigation environment and trusted forensic techniques, as well as allowing for the isolation of the database server from the network to prevent users from tampering with and/or capturing volatile and non-volatile data. Also, we determined that the PPPRC will have six investigation stages as shown in Fig 1, which are:

i. Notifying of Incident
ii. Incident Responding
iii. Identifying Source
iv. Verifying of Incident
v. Isolating Database Server
vi. Preparing Investigation Environment

The first investigation stage is Notifying of Incident. The DBA of the company notifies the higher management staff

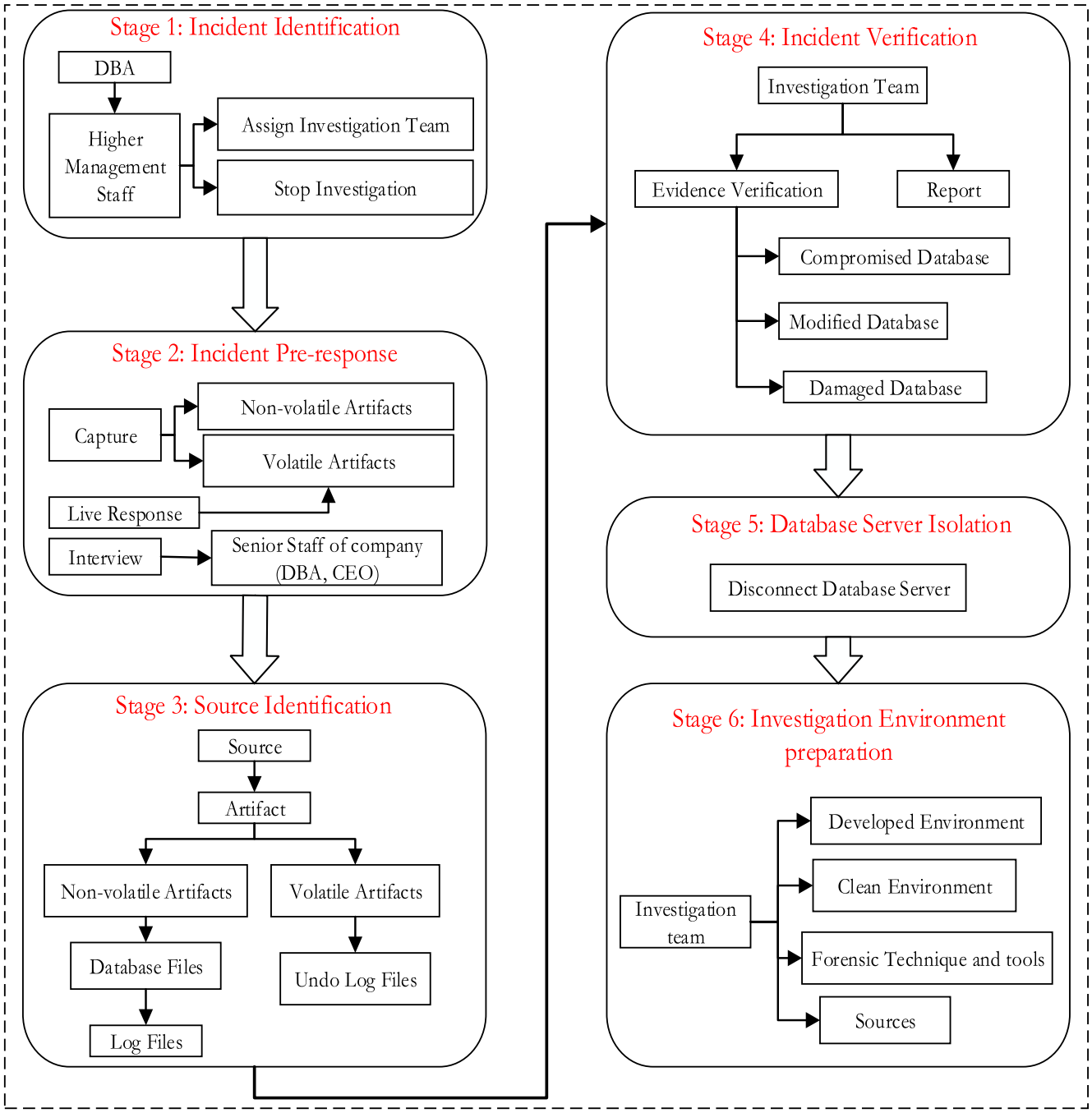

FIGURE 1. Planning, preparation and pre-incident responst category (PPPRC).

(e.g.: Chief Executive Officer (CEO), Chief Operating Officer (COO), Chief Security Officer (CSO)) about the database server incident [37]. In this case, the CEO of the company has two choices [10]: Either to assign an internal/external investigation team to investigate the database server incident, or stop the investigation [10]. The first choice: assign and authorize an internal/external investigation team. The investigation team performs the second stage of the PPPRC namely the Incident Responding Stage. The Incident Responding Stage is using for gathering incident details such as any information about incident events, parties involved thus far in the investigation, and the size and number of databases involved [8], [10]. The investigation team used trusted and cleaned forensic techniques to seize investigation sources [30], and gather the volatile artifacts [9], as well as gather valuable information through conducting interviews with staff [11]. Incident Responding Stage includes three concepts: Capture, Live Response, and Interview. In this way, the investigation team captures the investigation sources such as volatile and non-volatile artifacts [9]. Also, Live Response has an association

relationship with Volatile Artifact. Thus, the investigation team gathers valuable volatile data from Volatile Artifact. The last concept of the Incident Responding Stage is the Interview concept. The investigation team should conduct interviews with senior staff of companies such as the DBA and CEO [11]. The basic information such as information accounts, network ports, database servers, users, incident reports, logs, investigation procedures and policies may be gathered during the interviews [11]. Clearly, the Incident Responding stage allows the investigation team to illustrate the boundaries of an incident, and then identify the investigation sources. The third investigation stage is Identifying Source Stage that is used to identify specific investigation sources [9], [10], [33]. An investigation source includes several valuable volatile and non-volatile artifacts that hold the valued evidences.

Therefore, this stage includes evidence items that were seized and captured during the responding stage such as source, artifact, volatile artifact, nonvolatile artifact, database files, log files, and undo log files. The fourth investigation stage is Verifying of Incident Stage that allows the investigation team to check and verify the database incident [8], [10]. It consists of nine (9) concepts as discovered from the literature: Investigation Team, Incident, Modified Database, Destroyed Database, Compromised Database, Types Of Incident, Company, Report, and Decision. Therefore, the investigation team should determine what kind of incident (compromised, destroyed or modified) [18], the nature and the status of the incident. Then the investigation team submits detailed reports about the incident to company management [30]. Company management reviews the reports, and makes decisions: either to continue the investigation task, stop it or to disconnect the database server from the network [9], [34]. After verifying and determining the nature of the incident, the Isolating Database Server Stage is started.

The Isolating Database Server Stage is the fifth investigation stage that allows the investigation team to isolate/disconnect a suspect database server from the network to avoid more tampering [10], [7]. It consists of three concepts as discovered from the literature: Investigation Team, Database Server, and Database Management System. The isolating/disconnecting of the suspect database server does not mean a shutdown of the database [9], just isolating the users from the database management system [7], [34]. Finally, the investigation team should conduct the Preparing Investigation Environment Stage.

The Preparing Investigation Environment stage allows the investigation team to prepare the investigation environment to conduct a full investigation task [10]. The investigation environment includes six (6) concepts: Investigation Team, Forensic Workstation, Clean Environment, Found Environment, Forensic Technique, and Source. The investigation team prepares the trusted forensic workstation which includes the trusted forensic technique(forensic tools, and methods), and the investigation sources which were identified in the identifying stage.

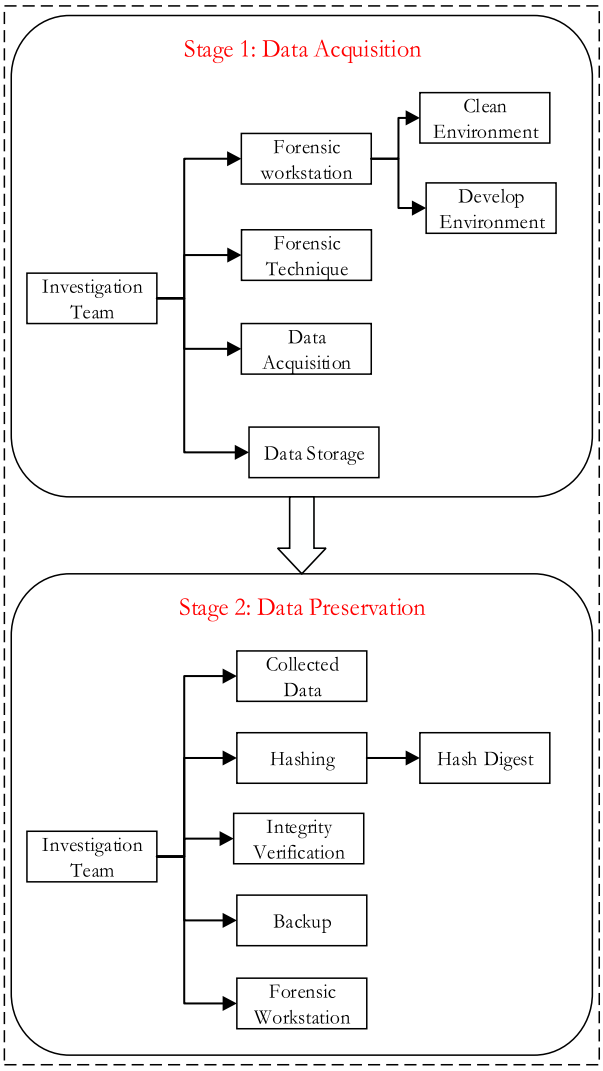

FIGURE 2. Acquisition and preservation category (APC).

B. CATEGORY B: ACQUISITION AND PRESERVATION CATEGORY (APC)

The second category is Category B, which contains 21 investigation processes that may use to collect and preserve volatile and non-volatile artifacts from the suspect database using trusted forensic techniques. We called this category Acquisition and Preservation Category (APC). we discovered the APC includes 2 investigation stages as shown in Fig 2 which are:

i. *Acquiring Data*
ii. *Preserving Data*

The Acquiring Data staged is used to gather/acquire data from a seized and captured investigation source that was identified in the identifying source stage [8], [9], [18], [30]. It consists of some concepts discovered from the literature to achieve this mission: Investigation Team, Report, Forensic Workstation, Clean Environment, Found Environment, Forensic

Technique, Data Acquisition, Source, and Data Collected. The Forensic Workstation concept includes trusted forensic techniques (forensic tools, and methods) to acquire Sources such as Volatile Artifact, and Nonvolatile Artifact. Investigation Team such as an investigator or examiner to achieve the Data Acquisition (Live Acquisition, Dead Acquisition, or Hybrid Acquisition) to acquire the volatile and non-volatile data from sources, which were seized and captured during the preparation stage. The output of this stage is the Data Collected. The Data Collected is data collected during the collection process that can be used for the analysis process. It includes many data relating to database activity, physical log files, and file database server. Furthermore, these data include evidences of what the intruder did and metadata regarding the intruder's activity [8], [9], [11], [18], [7], [29], [30], [33], [34], [38]. Therefore, the results of the Acquiring stage need to be preserved.

The Preserving Data stage is used to protect the integrity of data collected using hashing and backup methods [7], [31], and also to prevent any modification of collected data [10], [34]. The preserving data stage consists of Data Collected, Hashing, Integrity, Backup, Hashed Value, and Forensic Workstation concepts discovered from the literature. The Data Collected produced from the Acquiring data stage needs Hashing, and Backing up, to keep the integrity of data. Hashing is used to ensure that the database forensics techniques that were applied to hash the collected data have not changed the data. Also, it assures the reliability of transferred collected data between the source and the destination [9], [7], [38], [45]. Moreover, the backup concept provides an exact copy of data collected that may be used as a second copy when original data has been altered [7], [31], [33], [38], [45]. Therefore, the copy of the hashed collected data should be transferred to the forensic workstation through the secure channels to conduct reconstruction and analysis activities.

C. CATEGORY C: ANALYSIS AND RECONSTRUCTING CATEGORY (ARC)

The third category is Category C. It consists of 21 investigation processes that may be used for analysis of acquired data, activity reconstruction and data recovery using special forensic techniques to reveal who is tampering, when and where the tampering happened and how the tampering happened. We called this category Analysis and Reconstructing Category (ARC). It is worth noting that the logic of breach in a Database posit that the event has occurred and will thus require a reconstruction process, after the relevant data has been identified and acquired We determined the ARC has two investigation stages as shown in Figure 3 which are:

i. Examine Data Collected
ii. Analyse Data Collected
iii. Reconstruct Data Collected

The Examine Data Collected stage is used to ensure that data collected is authentic and has not been tampered with [7], [32], [34]. It consists of nine (9) concepts discovered from the

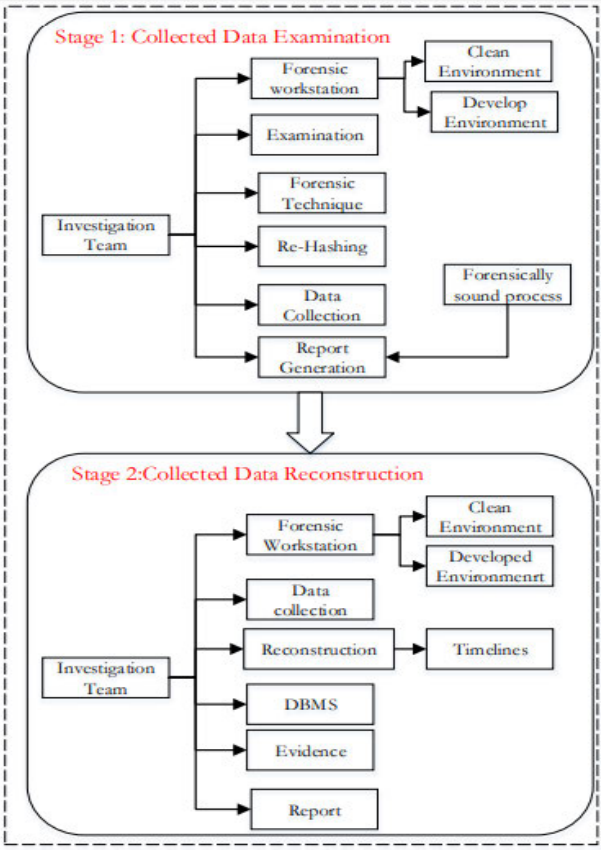

FIGURE 3. Analysis and reconstructing category (ARC).

literature: Investigation Team, Report, Forensic Technique, Examination, Data Collected, Forensic Workstation, Clean Environment, Found Environment, and Rehashing. Thus, the first mission of the investigation team is to examine the authenticity of data collected using appropriate forensic techniques.

However, if the collected data has been modified, the investigation team must bring another clean copy of the data collected from the originally collected data. The examination report is issued by the investigation team to document the steps and results of the examined data collected stage.

Next, the analysis allows the digital forensic experts to filter data acquired from a target device. Thus, the digital forensic examiner (In stage 2) will acquire data from the target, normalize the data, rehash the data again to ensure integrity, and store this data before the results are interpreted. According to [49], [50], such a process increases the potential of incident detection by generating forensic hypothesis that can be used to answer questions relative to security incidents. Answering these questions facilitates the reconstruction of the timeline of events

The Reconstruct Data Stage is used to reconstruct timeline events from collected volatile and non-volatile data which involves retracing past system, user database activity, past SQL execution history, stored procedures, and function execution [10], [18], [7], [33], [37]–[39]. It consists of nine (9) concepts as discovered from the literature: Forensic

Workstation, Reconstruction, Timeline, Data Collected, Investigation Team, Report, Forensic Technique, Database Management System, and Evidence. The investigation team, such as the examiner or analyzer, performs a reconstruction process using forensic techniques such as LogMiner [46], forensic algorithms [47], or Dragon [15]. The reconstruction process requires clean or existing DBMS and Data Collected to construct the Time Line. The timeline is a collection of digital events that have been recognized from the reconstruction process that will be used during analysis [8]. As an example of digital events that have been recognized: failed login events, successful login events, malicious database events that can be recognized and added to an examination timeline [33]. Furthermore, creating a timeline of events can assist an investigator to gain insight into the events that occurred and the people involved [39]. The Timeline concept has an association relationship with Forensic Technique, which may be used to search and filter the Timeline to offer the Evidence. Pieces of evidence are usually recognized in the database files that are recorded on hard drives and storage devices and media [32]. It is transmitted in binary form that may be relied upon in court [33]. It consists of who, why, when, what, how and where the malicious transactions were carried out [37]. Finally, the investigation team documents the whole reconstruction stage in several reports that should be submitted to the company and the court. Therefore, this study harmonized, categorized, and organized the redundant and overlapping DBFI investigation processes in specific and abstract categorizations. As asserted in [2], such processes provide a foundational basis for the development of a formalism which can be used in legal procedure. The scientific component of this process can further serve as a litmus test for suitability of such a harmonization.

Database forensics is a major consideration for both academics and practitioners. This study presented a systematic view of the overlapping processes in existing database investigation models. To achieve this, a research design approach was developed. The output of the design science process generated a generalized harmonized database forensic model that attempts to prevent the typical overlaps associated with existing database forensic models. This, therefore, provides a fundamental baseline for the evaluation of digital forensics processes in a manner that can survive legal scrutiny. Furthermore, the developed framework provides a substratum for the development of a database forensic ontology; a formalized process capable of aligning the forensic processes to a formal structure.

Future work includes the development of a formalized ontology, which can be integrated into any database investigation process. Leveraging ontology for formalism can potentially facilitate effective standardization and enhance the process of potential evidence admissibility.

V. CONCLUSION

The logic of categorizing the digital investigation process for A total of 40 DBFI process models were reviewed in this article. Process model researchers have used different approaches with different stages/phases and terminology. Most DBFI process models are specific and focus on specific RDBMS events, so they only provide low-level details. Furthermore, none of the studied DBFI process models can be called 'standardised' as each model has a different perspective. This paper contributes to the DBFI field by presenting a broad literature review that will assist field researchers in comprehending DBFI. This study studies all existing DBFI works, discuss the issues and drawbacks of the DBFI field, and suggest some solutions for the discovered limitations. The following are a few ideas for future works in the DBFI field: 1) the proposal of a generic DBFI process/model for the DBFI field; 2) the development of a semantic metamodeling language that structures, manages, organizes, shares, and reuses DBFI knowledge; and 3) the development a DBFI repository for the storage and retrieval of DBFI field knowledge.

REFERENCES

[1] N. Karie and H. Venter, "Resolving terminology heterogeneity in digital forensics using the Web," in *Information Warfare and Security*. Reading, U.K.: Academic Conferences and Publishing International Limited, 2013, p. 328.
[2] D. Ellison, A. R. Ikuesan, and H. Venter, "Description logics and axiom formation for a digital forensics ontology," in *Proc. Eur. Conf. Cyber Warfare Secur. Academic Conf. Int. Ltd.*, 2019, p. 742.
[3] A. R. Ikuesan and H. S. Venter, "Digital behavioral-fingerprint for user attribution in digital forensics: Are we there yet?" *Digit. Invest.*, vol. 30, pp. 73–89, Sep. 2019.
[4] A. Singh, A. R. Ikuesan, and H. S. Venter, "Digital forensic readiness framework for ransomware investigation," in *Proc. Int. Conf. Digit. Forensics Cyber Crime*. Midrand, South Africa: Springer, 2018, pp. 91–105.
[5] S. R. Selamat, R. Yusof, and S. Sahib, "Mapping process of digital forensic investigation framework," *Int. J. Comput. Sci. Netw. Secur.*, vol. 8, no. 10, pp. 163–169, 2008.
[6] Y. Yusoff, R. Ismail, and Z. Hassan, "Common phases of computer forensics investigation models," *Int. J. Comput. Sci. Inf. Technol.*, vol. 3, no. 3, pp. 17–31, Jun. 2011.
[7] D. Wong and K. Edwards, "System and method for investigating a data operation performed on a database," U. S. Patent 20 050 289 187 A1, Dec. 29, 2005.
[8] K. Fowler, G. C. F. A. Gold and M. MCSD, "A real world scenario of a SQL Server 2005 database forensics investigation," in *Information Security Reading Room Paper*. Bethesda, MD, USA: SANS Institute, 2007.
[9] D. Litchfield, "Oracle forensics—Part 4: Live response," NGSSoftware Insight Secur., New York, NY, USA, Tech. Rep., 2007.
[10] K. Fowler, *SQL Server Forenisc Analysis*. London, U.K.: Pearson 2008.
[11] N. Son, K.-G. Lee, S. J. Jeon, H. Chung, S. Lee, and C. Lee, "The method of database server detection and investigation in the enterprise environment," in *Proc. FTRA Int. Conf. Secure Trust Comput., Data Manage., Appl.* Berlin, Germany: Springer, 2011, pp. 164–171.
[12] H. Khanuja and S. S. Suratkar, "Role of metadata in forensic analysis of database attacks," in *Proc. IEEE Int. Advance Comput. Conf. (IACC)*, Feb. 2014, pp. 457–462.
[13] J. O. Ogutu and E. O. Abade, "A methodology to test the richness of forensic evidence of database storage engine," Analysis MySQL Update Operation InnoDB MyISAM Storage Engines, Univ. Nairobi, Nairobi, Kenya, Tech. Rep., 2016.
[14] A. Basu. (2006). *Forensic Tamper Detection in SQL Server*. [Online]. Available: http://www.sqlsecurity.com/chipsblog/archivedposts
[15] K. Pavlou, "Database forensics in the service of information accountability," NGSSoftware Insight Secur., New York, NY, USA, Tech. Rep., 2007.
[16] D. Litchfield, "Oracle forensics—Part 5: Finding evidence of data theft in the absence of auditing," in *NGSSoftware Insight Security Research (NISR)*. Sutton, U.K.: Next Generation Security Software Ltd, 2007.

[17] G. T. Lee, S. Lee, E. Tsomko, and S. Lee, "Discovering methodology and scenario to detect covert database system," in *Proc. Future Gener. Commun. Netw. (FGCN)*, Dec. 2007, pp. 130–135.
[18] O. M. Fasan and M. S. Olivier, "On dimensions of reconstruction in database forensics," in *Proc. WDFIA*, 2012, pp. 97–106.
[19] H. Beyers, M. S. Olivier, and G. P. Hancke, "Arguments and methods for database data model forensics," in *Proc. WDFIA*, 2012, pp. 139–149.
[20] J. Azemović and D. Mušić, "Efficient model for detection data and data scheme tempering with purpose of valid forensic analysis," in *Proc. Int. Conf. Comput. Eng. Appl. (ICCEA)*, Jun. 2009, pp. 1–8.
[21] J. Azemovic and D. Mušic, "Methods for efficient digital evidences collecting of business proceses and users activity in eLearning enviroments," in *Proc. Int. Conf. e-Educ., e-Bus., e-Manage. e-Learn.*, Jan. 2010, pp. 126–130.
[22] H. Beyers, M. Olivier, and G. Hancke, "Assembling metadata for database forensics," in *Proc. IFIP Int. Conf. Digit. Forensics* Berlin, Germany: Springer, 2011, pp. 89–99.
[23] D. Litchfield, "Oracle forensics—Part 2: Locating dropped objects," in *NGSSoftware Insight Security Research (NISR)*. Manchester, U.K.: Next Generation Security Software, 2007.
[24] A. C. Lawrence, "Forensic investigation of MySQL database management system," Univ. Cork, Cork, U.K., Tech. Rep., 2014.
[25] A. Al-Dhaqm, S. Abd Razak, S. H. Othman, A. Nagdi, and A. Ali, "A generic database forensic investigation process model," *Jurnal Teknologi*, vol. 78, nos. 6–11, Jun. 2016.
[26] A. Al-Dhaqm, S. Razak, S. H. Othman, K.-K.-R. Choo, W. B. Glisson, A. Ali, and M. Abrar, "CDBFIP: Common database forensic investigation processes for Internet of Things," *IEEE Access*, vol. 5, pp. 24401–24416, 2017.
[27] R. Bria, A. Retnowardhani, and D. N. Utama, "Five stages of database forensic analysis: A systematic literature review," in *Proc. Int. Conf. Inf. Manage. Technol. (ICIMTech)*, Sep. 2018, pp. 246–250.
[28] S. Kelly and R. Pohjonen, "Worst practices for domain-specific modeling," *IEEE Softw.*, vol. 26, no. 4, pp. 22–29, Jul. 2009.
[29] D. Lee, J. Choi, and S. Lee, "Database forensic investigation based on table relationship analysis techniques," in *Proc. 2nd Int. Conf. Comput. Sci. Appl.*, Dec. 2009, Art. no. 5404235.
[30] J. Choi, K. Choi, and S. Lee, "Evidence investigation methodologies for detecting financial fraud based on forensic accounting," in *Proc. 2nd Int. Conf. Comput. Sci. Appl.*, Dec. 2009, Art. no. 5404202.
[31] M. S. Olivier, "On metadata context in database forensics," *Digit. Invest.*, vol. 5, nos. 3–4, pp. 115–123, Mar. 2009.
[32] S. Tripathi and B. B. Meshram, "Digital evidence for database tamper detection," *J. Inf. Secur.*, vol. 3, no. 2, p. 113, 2012.
[33] H. K. Khanuja and D. Adane, "A framework for database forensic analysis," *Comput. Sci. Eng., Int. J.*, vol. 2, no. 3, pp. 27–41, 2012.
[34] R. Susaimanickam, *A Workflow to Support Forensic Database Analysis*. Murdoch WA, Australia: Murdoch Univ., 2012.
[35] H. K. Khanuja and D. D. S. Adane, "Forensic analysis of databases by combining multiple evidences," *Int. J. Comput. Technol.*, vol. 7, no. 3, pp. 654–663, Jun. 2013.
[36] H. Q. Beyers, "Database forensics: Investigating compromised database management systems," *Int. J. Comput. Technol.*, vol. 7, no. 3, pp. 654–663, 2013.
[37] P. Frühwirt, P. Kieseberg, K. Krombholz, and E. Weippl, "Towards a forensic-aware database solution: Using a secured database replication protocol and transaction management for digital investigations," *Digit. Invest.*, vol. 11, no. 4, pp. 336–348, Dec. 2014.
[38] O. M. Adedayo and M. S. Olivier, "Ideal log setting for database forensics reconstruction," *Digit. Invest.*, vol. 12, pp. 27–40, Mar. 2015.
[39] J. Wagner, A. Rasin, and J. Grier, "Database forensic analysis through internal structure carving," *Digit. Invest.*, vol. 14, pp. S106–S115, Aug. 2015.
[40] A. Al-Dhaqm, S. Abd Razak, S. H. Othman, A. Nagdi, and A. Ali, "A generic database forensic investigation process model," *Jurnal Teknologi*, vol. 78, nos. 6–11, Jun. 2016.
[41] M. F. Caro, D. P. Josyula, M. T. Cox, and J. A. Jiménez, "Design and validation of a metamodel for metacognition support in artificial intelligent systems," *Biologically Inspired Cognit. Archit.*, vol. 9, pp. 82–104, Jul. 2014.
[42] A. C. Bogen and D. A. Dampier, "Preparing for large-scale investigations with case domain modeling," in *Proc. DFRWS*, Aug. 2005, pp. 1–10.
[43] D. Litchfield, "Oracle forensics—Part 4: Live response," in *NGSSoftware Insight Security Research (NISR)*. Sutton, U.K.: Next Generation Security Software Ltd., 2007.
[44] J. Choi, K. Choi, and S. Lee, "Evidence investigation methodologies for detecting financial fraud based on forensic accounting," in *Proc. 2nd Int. Conf. Comput. Sci. Appl.*, Dec. 2009, Art. no. 5404202.
[45] H. Q. Beyers, "Database forensics: Investigating compromised database management systems," Sutton, U.K., Tech. Rep., Apr. 2019.
[46] P. M. Wright, "Oracle database forensics using LogMiner," in *Proc. Conf., SANS Inst.*, Jan. 2005, pp. 1–39.
[47] K. E. Pavlou and R. T. Snodgrass, "Forensic analysis of database tampering," *ACM Trans. Database Syst. (TODS)*, vol. 33, no. 4, p. 30, 2008.
[48] *Information Technology-Security Techniques-Incident Investigation Principles and Processes*, Standard ISO/IEC 27043, 2015. Accessed: May 2020. [Online]. Available: https://www.iso.org/standard/44407.html
[49] V. Kebande and H. Venter, "Towards a model for characterizing potential digital evidence in the cloud environment during digital forensic readiness process," in *Proc. ICCSM 3rd Int. Conf. Cloud Secur. Manage., ICCSM*, Oct. 2015, p. 151.
[50] V. R. Kebande and H. S. Venter, "Novel digital forensic readiness technique in the cloud environment," *Austral. J. Forensic Sci.*, vol. 50, no. 5, pp. 552–591, Sep. 2018.

ARAFAT AL-DHAQM (Member, IEEE) received the B.Sc. degree in information system from the University Technology of Iraq, and the M.Sc. degree (Hons.) in information security and the Ph.D. degree in computer science from University Technology Malaysia (UTM). He is currently working as a Postdoctoral Fellow with University Technology Malaysia (UTM). His doctoral research focused on solving the heterogeneity and ambiguity of the database forensic investigation field using a meta-modeling approach. His current research interests include digital forensics and cybersecurity.

SHUKOR ABD RAZAK (Member, IEEE) is currently an Associate Professor with Universiti Teknologi Malaysia. He is also actively conducts several types of research in digital forensic investigation, wireless sensor networks, and cloud computing. He is a author and coauthor of many journals and conference proceedings at national and international levels. His research interests include the security issues for mobile *ad hoc* networks, mobile IPv6, vehicular *ad hoc* networks, and network security.

DAVID A. DAMPIER (Senior Member, IEEE) received the Ph.D. degree in computer science from the Naval Postgraduate School. He is currently an Associate Dean of research with the College of Engineering and Computer Sciences, Marshall University. Prior to Marshall, he was the Chair of the Department of Information Systems and Cyber Security, The University of Texas at San Antonio, and the Founding Director of the Distributed Analytics and Security Institute, Mississippi State University. He has 70 peer-reviewed publications and ~$50M in external funding. Before academia, he spent 20 years as an Army Automation Officer. His research interests include cyber security, digital forensics, and applications of software engineering.

KIM-KWANG RAYMOND CHOO (Senior Member, IEEE) received the Ph.D. degree in information security from the Queensland University of Technology, Australia, in 2006. He currently holds the Cloud Technology Endowed Professorship at The University of Texas at San Antonio (UTSA). In 2015, he and his team won the Digital Forensics Research Challenge organized by the Germany's University of Erlangen–Nuremberg. He is also a Fellow of the Australian Computer Society. He was a recipient of the 2019 IEEE Technical Committee on Scalable Computing (TCSC) Award for Excellence in Scalable Computing (Middle Career Researcher), the 2018 UTSA College of Business Col. Jean Piccione, the Lt. Col. Philip Piccione Endowed Research Award for Tenured Faculty, an Outstanding Associate Editor of the IEEE Access, in 2018, the British Computer Society's 2019 Wilkes Award Runner-up, the 2019 EURASIP Journal on Wireless Communications and Networking (JWCN) Best Paper Award, the Korea Information Processing Society's Journal of Information Processing Systems (JIPS) Survey Paper Award (Gold) 2019, the IEEE Blockchain 2019 Outstanding Paper Award, the International Conference on Information Security and Cryptology (Inscrypt 2019) Best Student Paper Award, the IEEE TrustCom 2018 Best Paper Award, the ESORICS 2015 Best Research Paper Award, the 2014 Highly Commended Award by the Australia New Zealand Policing Advisory Agency, the Fulbright Scholarship, in 2009, the 2008 Australia Day Achievement Medallion, and the British Computer Society's Wilkes Award, in 2008. He is the Co-Chair of the IEEE Multimedia Communications Technical Committee's Digital Rights Management for Multimedia Interest Group.

KAMRAN SIDDIQUE (Member, IEEE) received the Ph.D. degree in computer engineering from Dongguk University, South Korea. He is currently an Assistant Professor with Xiamen University Malaysia. His research interests include cybersecurity, machine learning, and big data processing.

RICHARD ADEYEMI IKUESAN received the M.Sc. and Ph.D. degrees (Hons.) in computer science from Universiti Teknologi Malaysia. He is currently an Active Researcher pioneering a Digital Policing and Forensic Project for developing nations, using Nigeria and South Africa, as a hub for West Africa and Southern Africa, respectively. He is also an Assistant Professor with the Cyber Security Section of the IT Department, Community College of Qatar.

ABDULHADI ALQARNI is currently an Assistant Professor and the Chairperson of the Computer Science and Engineering Department, Jubail University College, Jubail Industrial City, Saudi Arabia. His research interests include internetworking switching and routing technologies, human–computer interaction, user experience, usable privacy and security, and data science.

VICTOR R. KEBANDE received the Ph.D. degree in computer science, majoring in information and computer security architectures and digital forensics, from the University of Pretoria, Hatfield, South Africa. He is currently a Postdoctoral Researcher with the Internet of Things and People (IoTaP) Center, Department of Computer Science and Media Technology, Malmö University, Sweden. His main research interests include cyber, information security and digital forensics in the area of the Internet of Things, (mainly IoT Security), digital forensics-incident response, cyber-physical system protection, critical infrastructure protection, cloud computing security, computer systems, distributed system security, threat hunting, modeling and cyber-security risk assessment, and blockchain technologies. He also serves as an Editorial Board member of *Forensic Science International: Reports* journal.