Acta Infologica
acin 2023, 7 (2), 396–408
DOI: 10.26650/acin.1142281

Review / Derleme

A Systematic Literature Review for New Technologies in IT Audit

Bilgi Teknolojileri Denetiminde Yeni Teknolojiler Üzerine Bir Sistematik Literatür Taraması

Nur Sena Tanrıverdi[1], Nazım Taşkın[2]

[1](PhD Candidate), Bogazici University, Faculty of Managerial Sciences, Department of Management Information Systems, Istanbul, Turkiye
[2](Assist. Prof.), Bogazici University, Faculty of Managerial Sciences, Department of Management Information Systems, Istanbul, Turkiye

Corresponding author : Nur Sena Tanrıverdi
E-mail : nur.tanriverdi@boun.edu.tr

ABSTRACT

Information technology (IT) audit focuses on auditing companies' IT systems and processes. The systems companies use are getting more complicated and better integrated. This means more data also needs to be audited. An IT audit often requires performing repetitive manual tasks, which makes IT audits more labor-intensive and costly. Current technological advancements have immense potential for improving an IT audit's performance, quality, and accuracy. Therefore, by leveraging advanced data processing and analysis technology, this workload can be lowered, allowing the auditing process to be performed effectively and efficiently with higher-quality outcomes. To achieve this objective, a systematic literature review (SLR) has been conducted to identify studies that use artificial intelligence (AI), machine learning (ML), predictive analytics, process mining, and natural language processing (NLP) techniques applied within IT auditing. Process mining is seen to have emerged as the most commonly used technique among the analyzed studies. The studies also reveal that combining techniques such as process mining and data mining, natural language processing, and machine learning enables effective and efficient audit processes by conducting continuous, automated, or online auditing work. The application of these new techniques in the examined studies are seen to generally provide solutions regarding the audit's testing stage. Overall, the study reveals a limited number of academic studies to have examined how these techniques are implemented into IT audits.

Keywords: Information technology audit, application audit, IT process audit, emerging technologies, systematic literature review

ÖZ

Bilgi teknolojisi (BT) denetimi, şirketlerin BT sistemlerini ve süreçlerini denetlemeye odaklanır. Bir şirkette kullanılan sistemler daha karmaşık ve entegre hale gelmektedir. Bu durum, denetlenmesi gereken verilerin artması ile sonuçlanır. BT denetimi genellikle manuel ve tekrarlanan görevlerin yapılmasını gerektirir, bu da BT denetimlerini daha emek yoğun ve maliyetli hale getirir. Mevcut teknolojik gelişmeler, BT denetim sürecinin performansını, kalitesini ve doğruluğunu iyileştirmek için büyük bir potansiyele sahiptir. Bu nedenle, gelişmiş veri işleme ve analiz teknolojisi kullanılarak bahsedilen iş yükü azaltılabilir ve denetim süreci daha kaliteli sonuçlarla etkin ve verimli bir şekilde gerçekleştirilebilir. Söz konusu amaca ulaşmak için, BT denetimi içinde uygulanan yapay zeka, makine öğrenimi, tahmine dayalı analitik, süreç madenciliği ve doğal dil işleme tekniklerini kullanan çalışmaları belirlemek için sistematik literatür taraması yapılmıştır. İncelenen çalışmalarda süreç madenciliğinin en çok kullanılan teknik olduğu görülmüştür. Çalışmalar ayrıca, süreç madenciliği ve veri madenciliği, doğal dil işleme ve makine öğrenimi gibi teknikleri birleştirmenin, sürekli, otomatik veya çevrimiçi denetim çalışması yürüterek etkin ve verimli denetim süreçlerine sahip olmayı mümkün kıldığını ortaya koydu. İncelenen çalışmalarda bu yeni tekniklerin uygulanması genellikle denetimin test aşamasına ilişkin çözümler sunmaktadır. Genel olarak çalışma, bu tekniklerin BT denetimlerine uygulanmasını inceleyen sınırlı sayıda akademik çalışma olduğunu ortaya koymaktadır.

Anahtar Kelimeler: Bilgi teknolojileri denetimi, bilgi teknolojileri genel kontrolleri, uygulama denetimi, BT süreç denetimi, gelişen teknolojiler, sistematik literatür taraması

Submitted : 12.07.2022
Revision Requested : 10.02.2023
Last Revision Received : 07.06.2023
Accepted : 10.07.2023
Published Online : 21.12.2023

1. INTRODUCTION

Information technology (IT) audit is an audit that focuses on companies' IT systems and processes. The audit process follows specific steps and tasks to determine critical issues and incompatible events within processes. At the end of the audit work, an audit opinion is provided to companies.

IT audit work is repetitive in nature (Manhanga, 2020), and performing repetitive manual tasks takes a serious amount of time and requires high labor costs. Moreover, the volume of data has increased as a result of the increasing number of integrated systems and digitalized business processes (Dzuranin & Mălăescu, 2016). Obtaining a high volume of data from systems and getting a sample of the data to audit are both challenges for an audit (Alexiou, 2019). Alongside sampling high volumes of data, another significant issue is examining all the data in a system.

However, systems have become capable of storing and processing high volumes of data, and analytical methods have also evolved. Therefore, current technology is also pushing changes in IT auditing. At the same time, new technologies can increase an IT audit's performance and lower the associated transaction costs while boosting the audit's quality and accuracy (Khan, Adi, & Hussain, 2021; Flores & Riquenes, 2020). Therefore, current top issues in auditing involve reducing the amount of time and costs required for audit work and increasing efficiency, effectiveness, and quality. These issues make the usage of new technologies in IT auditing crucial.

The Information Systems Audit and Control Association (ISACA, 2019a) conducted a survey among IT auditors about future predictions and directions of IT audits. This survey asked IT auditors about current technology usage in IT audit and predictions about technologies that will be used in the future. According to the survey results, 26% of IT auditors claimed their firms to use process mining, and 23% stated predictive analytics to be in use. In the report, 23% and 26% of the auditor participants respectively predicted that process mining and predictive analytics will still be used in the next 1-2 years. Additionally, 25% of the participants mentioned planning to use artificial intelligence (AI) within their audit work in the next 3-5 years, while 24% said they will use machine learning (ML) in the future. As explained above, IT audit work creates a substantial workload that is projected to grow even further in the future. However, this workload can be mitigated by using advanced data processing and analysis technologies, resulting in an efficient and effective auditing process with enhanced accuracy. Therefore, this literature review aims to identify studies that employ AI, ML, predictive analytics, process mining, and natural language processing (NLP) techniques. These identified techniques are currently being utilized or are anticipated to gain more prominence within the IT auditing domain. Their application facilitates more efficient and effective IT auditing, ultimately elevating the quality of audits. Additionally, this study is the first attempt to gather different approaches concerning the adoption of new technologies in IT auditing, an area limited by current techniques.

The primary aim of this study is to describe technologies that have already been integrated or that have the potential to be integrated into IT audits. By providing key insights from existing studies in the field of IT auditing, this review study aims to provide a comprehensive overview of the topic. With this aim, the study has created the following research questions:

RQ-1: What are the common approaches proposed for the use of new technologies/techniques in IT auditing?

RQ-2: In which step of IT auditing will the new technologies/techniques be used?

RQ-3: What data analytics algorithms are commonly used in IT auditing?

2. BACKGROUND INFORMATION

This section will provide background information about the key areas of this study, IT audits, and new technologies before applying the literature review methodology and presenting the related literature. After explaining a solid understanding of these key areas, the literature review methodology will be constructed and implemented.

2.1. IT Audit

IT auditing implies assessing and investigating all IT processes and the systems that process and store a company's critical data, such as financial data, customer data, and employees' personal data. An IT audit tests the defined control environment of a company and evaluates its compatibility with the company's policies (Carlin & Gallegos, 2007). An IT audit may also require an assessment of the organization's usage of IT in supporting the efficiency, effectiveness, and economics of its business processes (Carlin & Gallegos, 2007).

An IT audit is performed based on the collected audit evidence. ISACA (2019b) defined audit evidence as set of data employed to substantiate the audit assessment. Reports, documents, and system logs are audit evidence obtained based on the audit scope and period. IT auditors examine the audit evidence in detail and can conduct investigations directly on the systems in addition to these collected documents and logs. Information can be extracted from a system during auditing. Auditors also perform many inquiries in an audit on conducting and monitoring specific systems and processes with the people who are responsible in order to understand the processes. Although some of the audit evidence within IT audits are extracted from systems, IT auditors also perform manual audits. According to Mendez (2020), an IT auditor conducts the IT audit by following certain steps, such as planning, fieldwork or documentation, reporting, and follow up.

ISACA is a global professional association focused on IT governance, audits, risk, and compliance (ISACA, n.d.). The association creates a framework for IT governance, audits, and risk areas by collecting ideas from practitioners and academicians within the global network. They also publish reports and articles on their areas of activity.

An IT audit can be performed for different purposes. For example, an IT general controls audit aims to evaluate whether the controls of financial systems have been sufficiently set and are being efficiently run within the auditing period of a financial system and also evaluate whether financial systems are reliable (Barta, 2018; Chen, Hsu, & Hu, 2021). This audit is performed mainly under financial audits (Krieger, Drews, & Velte, 2021). Application checks can also be run within this type of audit; however, an IT application controls audit is subjected to evaluating the application itself (Barta, 2018; ISACA, 2019b). All IT components and systems are regulated and monitored under definitive process. Therefore, these processes (i.e., IT processes) are generally subject to an audit (ISACA, 2019b).

ISACA (2020) conducted a survey to provide a benchmark for IT audits. According to the frequencies of the answers the survey participants gave, the survey results show IT audits to be responsible for conducting IT general control audits, application audits, and IT process audits. Therefore, these audit types have been identified as the main focus of IT audits in this literature review study.

2.2. New Technologies

Based on the research ISACA (2019a) conducted, how IT audits are currently used, and what is expected to be used regarding IT audits in the future, this study has identified four technologies and techniques: AI, ML, predictive analytics, and process mining. These technologies and techniques mostly require structured data. However, a lot of unstructured data is also found to be audited as audit evidence in an IT audit. Therefore, natural language processing is also involved in the technologies on which this study focuses in order to cover the unstructured textual data analysis based on Schumann and Marx Gómez's (2021) approach.

3. RESEARCH METHODOLOGY

This paper uses a systematic literature review (SLR) methodology to comprehensively review new technology usage in the IT audit literature. SLR facilitates evidence-based guidelines for researchers and gives more rigorous results than ad-hoc literature review (Kitchenham et al., 2009; Tranfield, Denyer, & Smart, 2003). This article follows three stages within the SLR similar to how Tranfield et al. (2003) conducted their study: planning the SLR, conducting it, and reporting on it.

The planning stage should determine the needs of the review and develop the literature review protocol. The aim and needs of the SLR have already been explained in the Introduction. Additionally, the paper has adopted Schumann and Marx Gómez's (2021) structured literature review protocol due to the focus of this study resembling their focus on using natural language processing in internal auditing. The review protocol is given in Table 1. They conducted searches on six databases: Science Direct, Scopus, Web of Science, ACM Digital Library, IEEE Xplore Digital Library, and American Accounting Association. Databases with broader coverage than the others were accessed through the university library website. Thus, all databases were used to search papers except for the American Accounting Association database.

Table 1. *Review Protocol*

Unit of analysis	Journal articles
Type of analysis	Qualitative
Time period	2000-2021
Search fields	Title, Abstract, Keywords
Databases	Science Direct, Scopus, Web of Science, ACM Digital Library, and IEEE Xplore Digital Library
Total number of articles used in this study	8

The journal articles were evaluated for their effects and contributions to their fields (Hult, Reimann, & Schicke, 2009). In addition, the time interval as a search criterion was chosen based on Schumann and Marx Gómez (2021)'s study. Therefore, articles published between 2000-2021 were selected.

When conducting the review stage, the determined literature review protocol was executed. Keywords should be determined to select the relevant literature, and the keywords were selected based on concepts clarified in Section 2 (Background Information; see Table 2). Each concept has been expressed using different words and word groups in order to have the most generic terms as keywords written in different formats.

Table 2. *List of Keywords and References*

Perspective	Keywords
Technology	"artificial intelligence", "AI"
	"predictive analytics"
	"process mining"
	"machine learning", "ML"
	"natural language processing", "NLP"
Audit	"IT audit", "information technology audit"
	"IT general control audit", "information technology general control audit", "ITGC"
	"application audit"
	"IT process audit", "information technology process audit"

The search query given in Table 3 was run on the selected databases using the determined criteria.

Table 3. *Database Search Query*

In Title-Abstract-Keywords: ("artificial intelligence" **OR** "AI" **OR** "predictive analytics" **OR** "process mining" **OR** "machine learning" **OR** "ML" **OR** "natural language processing" **OR** "NLP") **AND** ("IT audit" **OR** "information technology audit" **OR** "IT general control audit" **OR** "information technology general control audit" **OR** "ITGC" **OR** "application audit" **OR** "IT process audit" **OR** "information technology process audit")

After the search was completed, 636 journal articles were collected from five databases. Of these articles, 317 are found on Science Direct, eight on Scopus, 289 on Web of Science, three on ACM Digital Library, and 19 on IEEE Xplore Digital Library (Figure 1). Only papers written in English were considered in this review, which then left 600 papers. The initial set was checked to identify duplicate papers. After removing duplicates, 530 unique journal articles were obtained in the initial data set.

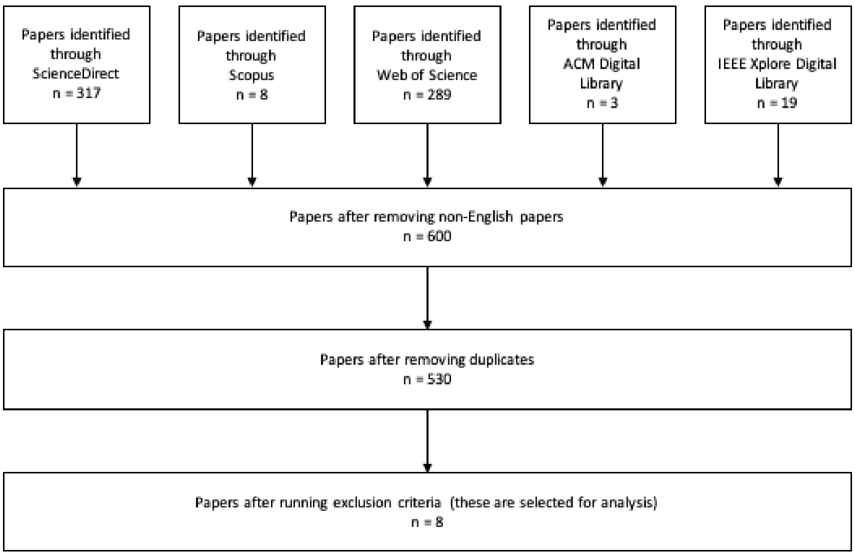

Figure 1. Flowchart of the literature Review process.

The aim of this literature review is to find studies that have applied such techniques as AI, ML, predictive analytics, process mining, and/or NLP with regard to IT auditing. This paper has identified studies that focus on tasks within any step of the IT auditing process. Therefore, the exclusion criteria given in Table 4 were formed and applied to the initial set of studies. The number of papers excluded from the initial set after applying the criteria are given in Table 4.

Applying the first exclusion criterion (C1) excluded 407 papers for containing only the searched keywords without the context of corporate auditing. The second exclusion criterion (C2) eliminated papers related to auditing that did not refer to conducting an audit or that did not contain such things as auditing tasks or auditing methodologies. In addition, papers that focused on the auditability of a proposed system, model, or new technology without giving an auditing approach or methodology were eliminated. This second criterion eliminated another 25 papers. The third exclusion criterion (C3) was run to eliminate papers that were related to auditing but that only handled the topic at a conceptual level. Based on this third exclusion criterion, another 16 papers were excluded. The main focuses of these eliminated papers involved such topics as listing the advantages of new approaches in auditing, their acceptance factors, auditors' qualifications, and changes in the audit profession.

After running the exclusion criteria C1, C2, and C3, the remaining papers are seen to include tasks and approaches within the auditing process. Starting from this point, the papers were analyzed with a greater focus on IT auditing. However, the remaining 74 papers did not include any studies implementing new techniques and technologies into application auditing, IT general controls (ITGC), IT process audit types in IT auditing, or other types of IT auditing.

Meanwhile, the final set of articles had some papers that focused on providing a method to audiences rather than just applying methods and techniques to specific auditing subjects. Jans and Hosseinpour (2019) emphasized the independence of the method from any specific auditing subject. When considering the potential applicability of the method papers to other audit subjects, the remaining set of articles was decided to be evaluated accordingly. Therefore, a fourth exclusion criterion was developed with this aim. According to the fourth exclusion criterion (C4), papers were excluded that provided specific solutions for certain audit subject data (e.g., tax, finance, fraud), because applying these solutions to IT auditing was unsuitable or evaluating their applicability to the IT audit area was ineffective. Meanwhile, papers related to business process auditing or internal auditing were not excluded because they can be applied to the field of IT audits. This criterion eliminated another 57 papers.

The fifth exclusion criterion (C5) filtered out an additional nine papers. These papers had no applications focused on AI, NLP, ML, predictive analytics, or process mining, which are the new technologies that have been determined in this research. Concepts covering such technologies as advanced data analytics or big data analytics have been included. For example, many papers provided blockchain applications within auditing as a new approach. The final exclusion criterion (C6) eliminated four papers that were inaccessible through the university library's access rights. Eight papers remained from the initial set after applying these exclusion criteria.

The criteria C1, C2, and C3 were adapted from Schumann and Marx Gómez's (2021) study, whereas the criteria C4, C5, and C6 were created to capture research suitable for this study. The exclusion criteria and the number of excluded papers are given in Table 4.

Table 4. *The Exclusion Criteria and Number of Excluded Papers*

ID	Exclusion Criteria	Number of Excluded Papers	ID	Exclusion Criteria	Number of Excluded Papers
C1	No actual auditing reference	407	C4	No approaches that are adaptable to the IT audit engagement process	57
C2	No task within the audit engagement process	25	C5	No reference to the relevant technology	9
C3	No technical approach	16	C6	Article is inaccessible	8

Eight papers were determined for the analysis after applying the exclusion criteria. The papers were then evaluated based on the categories given in Table 5. These categories were created deductively. Eight categories (i.e., audit type, IT audit stage, new techniques/technologies, algorithm, data type, methodology and framework, validation, and evaluation) were created by following Schumann and Marx Gómez's (2021) approach. The full text of the eight papers were qualitatively analyzed based on these categories in order to fulfill the research questions. Thus, planning and conducting the SLR stages have been performed so far.

Table 5. *Categories for Literature Analysis*

Category	Description of Category
Audit Type	Type of audit (application audit, IT process audit, ITGC, others)
IT Audit Stage	Stage of IT audit (planning, fieldwork/documentation, report/follow up)
New Techniques/Technologies	Determined new techniques/technologies (AI, NLP, ML, predictive analytics, process mining)
Algorithm	Applied data analytics algorithm (e.g. clustering, support vector machines, etc.)
Data Type	Type of inputted data of the model (unstructured, structured)
Methodology and Framework	Whether methodology or framework is provided in the paper
Validation	Whether validation is performed in the paper
Evaluation	Whether performance of techniques is evaluated in the paper

4. RESULTS

This section reports on the literature review in stages similar to those Tranfield et al. (2003) suggested for a literature review. The following paragraphs and sections give descriptive and thematic analyses of the articles.

Among the eight articles in the final set, the earliest publication occurred in 2008, while the most recent articles were published in 2021 (Table 6). As seen in Table 6, three recent papers published in 2020 and 2021 employed AI, ML, and natural language processing. While four of the papers published before 2020 employed process mining, one employed data mining.

Table 6. *The Papers' Publication Year, Audit Type, and New Techniques/Technologies*

Reference	Year	Audit Type	New Techniques/Technologies
Rozinat & van der Aalst (2008)	2008	Other	Process mining
Caron, Vanthienen, & Baesens (2013)	2013	Other	Process mining
Kuna, Gartia-Martinez, & Villatoro (2014)	2014	Other	Other (Data mining)
Zerbino, Aloini, Dulmin, & Mininno (2018)	2018	Other	Process mining
Jans & Hosseinpour (2019)	2019	Other	Process mining & Other (Data mining)
Yesmin & Carter (2020)	2020	Other	ML
Khan et al. (2021)	2021	Other	ML & NLP
Chen et al. (2021)	2021	Other	AI

The following subsections provide thematic analyses of the eight papers. These eight papers have been analyzed based on the eight categories given in Table 5. The literature analysis results are presented based on categories and in separate subsections titled the same as the category names except for the categories of validation and evaluation, which have been grouped together under the subsection titled Technique Validation and Evaluation.

4.1. Audit Types

Analysis of the papers shows none of the articles to focus on implementing new techniques into IT audits. The aim was to find the most frequently conducted IT audit types in the literature identified by the keywords of IT process audit, application audit, and ITGC. Other IT audit types and the specified IT audit types were not found in the literature. However, audit types other than IT audits (e.g., internal audit, business process audit0 were found in the literature and identified in Table 6 as "Other;" these were then analyzed to determine their audit subjects. As a result, business process audits are dominant in the final set.

In some papers, the authors named the subject of the audit work as an internal audit (Chen et al., 2021; Jans &

Hosseinpour, 2019), information systems audit (Kuna et al., 2014; Zerbino et al., 2018), or business process audit (Caron et al., 2013; Khan et al., 2021; Rozinat & van der Aals, 2008). In these articles, internal audits and information systems refer to business process audits. These articles focused on different business processes such as innovation management (Khan et al., 2021), claim-handling processes (Caron et al., 2013) and freight export port processes (Zerbino et al., 2018).

One article focused on a privacy audit (Yesmin & Carter, 2020). Privacy audits can be evaluated as a business process audit. However, the article's auditing process involved technical auditing, not business process auditing. The methods and techniques studied in these papers, which had been predominantly subjected to business process audits, were evaluated, and these methods and techniques are considered suitable for adapting to IT auditing.

4.2. IT Audit Stages

The papers were analyzed based on the audit stages on which they were focused. Except for Chen et al.'s (2021) study, all papers focused on the testing stage. Chen et al. (2021) focused on the planning stage of an IT audit with the aim of identifying critical risks in an audited company as planning for an internal control audit. They developed a hybrid decision-making model that involved data exploration.

4.3. New Techniques

Several techniques were used in the papers, including process mining, ML, NLP, and AI. Most papers developed an approach with process mining (Caron et al., 2013; Jans & Hosseinpour, 2019; Rozinat & van der Aalst, 2008; Zerbino et al., 2018). Jans and Hosseinpour (2019) combined process mining and data mining. AI (Chen et al., 2021), ML techniques (Khan et al., 2021; Yesmin & Carter, 2020), NLP techniques (Khan et al., 2021), and data mining techniques (Kuna et al., 2014) were employed within the testing stages of audits. The testing stage constitutes the main work and mostly entails repetitive labor-intensive work.

The four process mining articles (Caron et al., 2013; Jans & Hosseinpour, 2019; Rozinat & van der Aalst, 2008; Zerbino et al., 2018) were method articles that developed a method using process mining. These papers employed conformance checking algorithms to determine the derivations of processes from as-is processes. The techniques performed in these papers also correspond to the testing stage.

Kuna et al. (2014) employed data mining techniques to detect anomalous actions within processes of an academic management system, purchase management system, and inventory control system. Yesmin and Carter (2020) employed supervised ML algorithms in order to determine inappropriate access to a hospital system. Khan et al. (2021) applied NLP and ML algorithms to text-based evidence to automatically review innovation management systems' compliance with the ISO56002 innovation management system standard. Therefore, anomalous data detection, inappropriate access identification, and compliance check of text-based audit evidence are the techniques that were performed within the testing step.

4.4. Algorithms

The studies employed different data analytics algorithms and combinations of algorithms. Because four process mining articles are found in the final set, the common approach among these studies can be more clearly identified. Conformance checking algorithms specially developed for compliance analysis were commonly used in the process mining articles (Caron et al., 2013; Jans & Hosseinpour, 2019; Rozinat & van der Aalst, 2008; Zerbino et al., 2018). Process discovery algorithms (e.g., Alpha algorithm, heuristic miner, fuzzy miner) are first used in process mining algorithms to determine the executed processes in the event log. A conformance checking algorithm is then used to test the compliance of the process executions to an as-is process model; this was determined using the process discovery algorithms.

Zerbino et al. (2018) analyzed data in two steps. Before conducting compliance tests, they first created a control-flow model. The control flow model was developed with process discovery techniques. They used a fuzzy miner algorithm for the process discovery. Once they had the control-flow model, they employed a conformance checking algorithm to identify non-conformances within the process. Rozinat and van der Aalst (2008) employed two different approaches to the conformance checking algorithm in their method that overcomes specific problems regarding log replay analysis. The first algorithm solves the problem of not being fired from an invisible task due to no related log event. The second algorithm was created to solve duplicate tasks resulting from a model task and log event mapping. Jans and Hosseinpour (2019) presented a different approach. They proposed applying both process mining and data mining algorithms within their framework. For their first step, they used the most suitable process discovery algorithm from among all the process discovery algorithms (e.g., alpha algorithm, heuristic miner, fuzzy miner, genetic miner, inductive miner) to identify executed processes. Then they ran a compliance checking algorithm to determine deviations in the process execution compared to the as-is process. After completing the process mining step and determining the deviations, they grouped the deviations using data mining classification algorithms. Their study combined two analytic approaches, while Caron et al. (2013) did not specify the algorithm in their study.

Additionally, no specific algorithm was seen to have been frequently used in the articles that use ML, NLP, AI, or data mining algorithms. Each article employed a different algorithm. However, two articles (Kuna et al., 2014; Yesmin & Carter, 2020) used algorithms to classify identified deficiencies. Firstly, Yesmin and Carter (2020) implemented supervised ML algorithms to proactively identify unauthorized employee access to the hospital system. Their approach also enabled the categorization of detected access into three groups: appropriate access, unexplained access, and flagged access. They also provide customized explanations for detected access. Kuna et al.'s (2014) study implemented various data mining algorithms to determine anomalous data in audit logs. They first employed LOF and DBSCAN algorithms for outlier detection within audit logs and then applied the C4.5, Bayesian Network (BN), and PART algorithms to classify the detected outliers.

Khan et al.'s (2021) study is the only one to have analyzed textual data while employing NLP. They applied NLP techniques such as lemmatization and stop-word removal for data cleansing, the clustering method for closeness and associations between words, and word frequency for visualizing textual data.

Meanwhile, Chen et al. (2021) applied a hybrid model that included various algorithms for identifying risks in internal auditing. They mentioned their integrated hybrid model to contain rule generation techniques based on AI algorithms such as support vector machines, ensemble learning, decision trees, and multiple rule-based decision-making techniques. However, they did not explain how they implemented the AI algorithms.

4.5. Methodology and Framework

This category presents the methodologies and frameworks provided in the articles. The authors provided their methodology, framework, or only the steps they followed while implementing the algorithm in their articles (Caron et al., 2013; Jans & Hosseinpour, 2019; Kuna et al., 2014; Rozinat & van der Aalst, 2008; Zerbino et al., 2018). Some authors developed an audit tool in their study (Khan et al., 2021; Yesmin & Carter, 2020). Only one study (Chen et al., 2021) did not explain their methodology.

Zerbino et al. (2018) proposed a process mining-enabled methodology for an information systems audit containing five steps: justification and planning, data extraction, control-flow model construction, model enrichment, and conformance checking. They used their methodology in an attempt to find a solution for ineffectiveness at detecting non-conformances, frauds, and abuses. Caron et al. (2013) proposed a technique for compliance checking due to the inadequacy of existing techniques in process discovery and visualization, conformance checking and delta analysis, and logic-based property verification with regard to compliance checking. Their proposed technique involved comprehensive rule-based compliance checking using process mining. The technique also contained the architecture for compliance checking, two dimensions of business rule taxonomy, rule restriction, formal specification of rule patterns, and definition of the scope of rules. Jans and Hosseinpour (2019) provided a transactional verification framework for continuous auditing. Their framework contained six phases from building an event log to identifying deviations in transactions by combining process mining and data mining techniques. Rozinat and van der Aalst (2008) proposed a conformance checker technique to extend current techniques. Their conformance checker is an incremental technique for validating compliance in a process model and an event log. The conformance checker provides new metrics with precise definitions on how to implement these metrics. Meanwhile, Kuna et al. (2014) presented no framework for the method they applied. However, they did provide their procedure clearly, including how to read data from the database, preprocess the data, apply outlier detection rules, and classify detected outliers.

Yesmin and Carter (2020) provide neither a defined methodology or framework for applying their technique. Nevertheless, they built an automatic privacy auditing tool, including their solution. They also presented a framework for evaluating the outputs of the employed technique. Khan et al. (2021) developed an AI-based audit tool with a fuzzy front-end of the ISO56002 Innovation Management System Standard. They also presented no framework but did provide their procedure, which involved the specific tasks of preprocessing input data sets, cluster mapping and visualizing the standard, visualizing the standard's cluster map with audit transcripts, normalizing Map 2, and automated scoring by creating a reverse map.

4.6. Validation and Evaluation of the Techniques

This section analyzes whether the studies included any validation or performance evaluation for their proposed models. The articles had different approaches for validating and evaluating their techniques. First, Zerbino et al. (2018) applied the proposed process mining-enabled methodology to the freight export port process managed by a port community system (PCS). They did not perform a numerical evaluation of the technique's performance but did verbally assess the technique, mentioning the drawbacks of their technique in the paper. These drawbacks are related to being familiar with the process and the information systems from which the input data for the process mining methods come with regard to determining deviations in processes.

Two application examples were provided in Rozinat and van der Aalst's (2008) study. The first was applied to the administrative processes of a municipality in the Netherlands. The other application analyzed the conformance of web service behavior. The technique provided in Rozinat and van der Aalst's study contained metrics for evaluating conformance checks, and they also approached how to evaluate the metrics.

Yesmin and Carter's (2020) approach was an application of a method rather than the creation of one. As such, they did not require validation for any method. However, they demonstrated the benefit of their method's results. In order to accomplish this, they provided a validation framework that included various scenarios about user behavior on the privacy auditing tool they had created. All scenarios were created from different groups of employees' perspectives and corrected by experts. Their framework enabled them to test their results with different use cases. Yesmin and Carter's evaluated how well their model determined categories and how it had evolved itself to make more accurate classifications. They explained that they had achieved a statistically significant learning curve as a result.

Kuna et al. (2014) implied the solution to three different systems: the academic management system of a university, the purchase management system of a local government, and the inventory control system of a wholesaler. They also provided an evaluation that compared false positives with the efficacy values of the results from the three applications. Moreover, to demonstrate their developed method's effectiveness and efficiency, Khan et al. (2021) compared the compliance scores of audits performed by the developed AI tool and manually conducted audits of five organizations. Khan et al. evaluated the efficiency and effectiveness metrics to compare the automated and manual audits.

Meanwhile, Caron et al. (2013) mentioned testing the proposed technique in claims-handling processes in the insurance industry. However, they did not provide any evaluation results. Additionally, Jans and Hosseinpour (2019) and Chen et al. (2021) neither evaluated nor conducted any validation for their proposed techniques.

4.7. Data Type

The structure and source of input data varied in the articles based on the employed techniques. Apart from Khan et al.'s (2021) study, all the others (Caron et al., 2013; Chen et al., 2021; Jans & Hosseinpour, 2019; Kuna et al., 2014; Rozinat & van der Aalst, 2008; Yesmin & Carter, 2020; Zerbino et al., 2018) used structured data.

Khan et al. (2021) used unstructured data, which was also text-based. They also used two different sets of data as inputs for the audit. The first were transcripts from the interviews about the conducted innovation management process in organizations. The other were the requirements for fuzzy front-end with regard to the ISO56002 Innovation Management System Standard in text format.

The articles that employed process mining as a technique used structured data. These data included event log data belonging to the executed process. The fundamental attributes of event log data are case ID as a unique identifier, event name as activity, and timestamp defining when the event was executed. Process mining studies in the final set used event log data belonging to different processes such as the administrative processes of a municipality in the Netherlands and processes on a web service (Rozinat & van der Aalst, 2008), claim-handling process (Caron et al., 2013), and freight export port process (Zerbino et al., 2018). Event log data were generally obtained from the system on which the processes run and on which event logs are simultaneously stored.

The other papers that employed ML and AI used structured data in their analysis (Chen et al., 2021; Kuna et al., 2014; Yesmin & Carter, 2020). Yesmin and Carter (2020) collected user and patient data from different systems such as human resources and scheduling software. Kuna et al. (2014) used audit logs from three different systems: academic management system, purchase management system, and inventory control system. Chen et al.'s (2021) data sources were the most varied among all the papers. They collected data using a questionnaire for risk identification, in which the participants were domain experts.

5. DISCUSSION

This study has performed an SLR to present an overview of papers that had implemented new technologies and techniques (i.e., AI, ML, NLP, predictive analytics, and process mining) into IT audits. Even if some papers had IT audits as a subject, no paper apparently focused on using new technology and techniques, especially in IT audits. However, the literature had some papers that provided a method for applying certain techniques to related audit areas, such as business process and information systems audits. Techniques applied in similar areas have the potential to be implemented in IT audits. Therefore, this study aimed at presenting an overview of these papers. Additionally, due to the study being the first one in the context of IT auditing, no earlier studies conducted in the literature could be referred to in this context.

The first finding from this literature review study is the lack of academic studies focused on implementing techniques into IT audits. In the research conducted among IT auditors by ISACA (2019a), however, 26% of the IT auditors claimed to currently use process mining and 23% to use predictive analytics. As such, the gap between practice and academic studies in terms of the technologies used in IT audit can be revealed.

As was given in the results, eight papers were identified in total. These papers included techniques adaptable to IT audits. One paper each contained the following techniques: artificial intelligence (Chen et al., 2021), machine learning (Yesmin & Carter, 2020), natural language processing (Khan et al., 2021), and data mining (Kuna et al., 2014). Meanwhile, four papers employed process mining (Caron et al., 2013; Jans & Hosseinpour, 2019; Rozinat & van der Aalst, 2008; Zerbino et al., 2018). Just as the IT auditors had mentioned their frequent usage of process mining in ISACA's (2019a) study, the most studied technique was process mining in the studies analyzed here. These papers also employed a combination of these techniques, such as process mining with data mining (Jans & Hosseinpour; 2019) or NLP with ML (Khan et al., 2021).

Seven papers (Caron et al., 2013; Jans & Hosseinpour, 2019; Khan et al., 2021; Kuna et al., 2014; Rozinat & van der Aalst, 2008; Yesmin & Carter, 2020; Zerbino et al., 2018) provided a methodology, framework, or procedure to explain what they had performed or proposed to apply as a technique. These papers are helpful for practitioners and academicians. Most of the papers had validated their proposed techniques with either actual data (Caron et al., 2013; Khan et al., 2021; Rozinat & van der Aalst, 2008; Yesmin & Carter, 2020; Zerbino et al., 2018) or artificial data (Kuna et al., 2014; Rozinat & van der Aalst, 2008). However, the authors from two studies did not validate their proposed methods (Chen et al., 2021; Jans & Hosseinpour, 2019). Additionally, providing performance evaluations in a paper, as had occurred in five of them (Khan et al., 2021; Kuna et al., 2014; Rozinat & van der Aalst, 2008; Yesmin & Carter, 2020; Zerbino et al., 2018), makes assessing the suggested technique more understandable for audiences.

From the audit perspective, seven papers (Caron et al., 2013; Jans & Hosseinpour, 2019; Khan et al., 2021; Kuna et al., 2014; Rozinat & van der Aalst, 2008; Yesmin & Carter, 2020; Zerbino et al., 2018) out of eight had clearly provided a solution for the testing stage of the audit. Only one study (Chen et al., 2021) handled the planning stage. However, considerable potential is found for applying different techniques to other audit stages. Whether the input data of the technique is structured or unstructured changed based on the technique. The input data for the techniques corresponded to the audit evidence. Thus, any kind of audit evidence can be used in the various techniques.

This paper can support researchers studying new technologies in IT auditing by providing insights. The results from this study may also help audit teams when they implement and develop new technological solutions for their audit processes. In addition, the main reasons are understood as to why studies apply new audit techniques. According to the authors, these reasons are to provide a continuous audit approach (Caron et al., 2013; Jans & Hosseinpour, 2019), to automate the audit work (Khan et al., 2021; Kuna et al., 2014; Zerbino et al., 2018), and to perform online or real-time audits (Zerbino et al., 2018) in organizations. These reasons justify the need to apply new techniques that add effectiveness and efficiency to the audit processes of today's increased audit workload. These reasons also support the aim of this literature review work. The studies can be concluded to have developed new approaches, mainly process mining, data mining, ML, and NLP techniques or some combination of these techniques and to have achieved effective and efficient audits by conducting continuous, automated, or online audit works. Meanwhile, this study has identified the gap between practice and academic studies in the field of IT audits. Closing this gap would expand the IT audit literature and also provide an example of how the related technologies are used. Both academicians and practitioners can benefit from closing the gap. One method developed by any company specific to their needs can be transformed into an academic study by providing a performance evaluation or even the insights of the involved parties who use these new approaches or have develop these techniques to understand the adaptation process through their experiences.

As mentioned before, after not finding studies related to implementing new technologies into IT audits, the articles that used these technologies for other audit areas were examined according to their applicability to IT audits. The existing literature may have more studies with the potential for use in IT audits. These articles were not fully covered in this study. Therefore, future systematic literature review studies can expand their search using additional keywords that can be selected according to the relevance of the area to IT audits.

Additionally, future literature review studies could specify research questions to specific IT domains or processes. Specific evidence types could also be determined, such as third-party contracts, previous years' audit reports, deployment logs of applications, or request records from request management systems.

Moreover, conference proceedings can also be evaluated. The current study could be enhanced with other emerging technologies such as blockchain and big data analytics. New technologies could also be enhanced with data mining because of its relevance.

This study also encountered other limitations with regard to the searched databases. Conceptualization was performed based on the articles published in ISACA's database; however, any advanced searches were not conducted on this database, because ISACA's publication database does not provide an advanced search function that would enable a systematic search, unlike the other databases. A review of the ISACA database would be beneficial.

6. FUTURE STUDIES AND RECOMMENDATIONS for IT AUDITS

Studies on IT auditing enriched with AI, ML, NLP, predictive analytics, and process mining are not available in the literature. Studies and real-life applications should focus on having continuous and automated auditing while closing the gap between practice and academic studies in the IT audit area.

Based on the reviewed literature, this study suggests that, in order to emphasize the value of implementing new techniques into the auditing process, validation and evaluation of proposed methods should occur, as many papers did to show the applicability and effectiveness of their methods. In addition, the method, framework, or steps a proposed technique follows should be given to enlighten people who are considering adopting the technique. Providing information regarding input data type (i.e., audit evidence) is also significant in clarifying technique details.

IT processes subject to audits such as access management, project management, or processes within the software development life cycle are suitable for assessment using process mining. A log of users' activities within these processes can be extracted from related systems. An IT process that starts at one system and ends up at another can be also analyzed with process mining techniques. Process mining can also be combined with machine learning or data mining techniques, such as the ones Jans and Hosseinpour (2019) performed for more sophisticated evaluations of IT audits.

As Khan et al. (2021) proposed, an NLP technique can be employed for conducting IT audits based on standards, frameworks, or regulations such as ISO27001, Cobit 4.1, SOX, or SSAE.

Most of the studies had provided a solution for the testing stage of an IT audit. As such, the potential could exist for focusing on the automatization of areas in the planning, reporting, and follow-up stages of IT audits rather than just focusing on the testing stage.

Additional suggestions have been intuitively made as follows. NLP techniques, especially in the planning stage, could be employed because textual data are generally used in that stage for determining the scope of audits and controls. For example, summarizing the previous year's audit reports, being one of the inputs of the planning stage, can be performed using NLP techniques. An audit's scope, risk, and controls can be identified for auditing an IT service provider using third-party agreements belonging to the provider with regard to outsourced IT services, and NLP can be used for this analysis. While concluding an audit, machine learning algorithms can also help determine the classification of deficiencies and create an audit opinion based on the audit results.

Peer Review: Externally peer-reviewed.

Author Contributions: Conception/Design of Study- N.S.T.; N.T.; Data Acquisition- N.S.T.; Data Analysis/Interpretation- N.S.T.; N.T.; Drafting Manuscript- N.S.T.; Critical Revision of Manuscript- N.S.T.; N.T.; Final Approval and Accountability- N.S.T.; N.T.

Conflict of Interest: The authors have no conflict of interest to declare.

Grant Support: The authors declared that this study has received no financial support.

ORCID IDs of the authors / Yazarların ORCID ID'leri

Nur Sena Tanrıverdi 0000-0002-4728-988X

Nazım Taşkın 0000-0002-5327-9012

REFERENCES

Alexiou, S. (2019). Trends, challenges and strategies for effective audit in a rapidly changing landscape. *ISACA Journal, 6*. Retrieved from https://www.isaca.org/resources/isaca-journal/issues/2019/volume-6/trends-challenges-and-strategies-for-effective-audit-in-a-rapidly-changing-landscape

Barta, G. (2018). The increasing role of IT auditors in financial audit: risks and intelligent answers. *Business, Management and Education, 16*(1), 81-93. https://doi.org/10.3846/bme.2018.2142

Carlin, A., & Gallegos, F. (2007). IT audit: A critical business process. *Computer, 40*(7), 87-89. https://doi.org/10.1109/MC.2007.246

Caron, F., Vanthienen, J., & Baesens, B. (2013). Comprehensive rule-based compliance checking and risk management with process mining. *Decision Support Systems, 54*(3), 1357-1369. https://doi.org/10.1016/j.dss.2012.12.012

Chen, F. H., Hsu, M. F., & Hu, K. H. (2021). Enterprise's internal control for knowledge discovery in a big data environment by an integrated hybrid model. *Information Technology and Management, 23*, 213-231. https://doi.org/10.1007/s10799-021-00342-8

Dzuranin, A. C., & Mălăescu, I. (2016). The current state and future direction of IT audit: challenges and opportunities. *Journal of Information Systems, 30*(1), 7-20. https://doi.org/10.2308/isys-51315

Flores, I. G., & Riquenes, J. R. (2020). Audit 2.0, a perspective for its execution in the business environment using process mining techniques. *Vivat Academia, 23*(150), 47-57. http://doi.org/10.15178/va.2020.150.47-57

Hult, G. T. M., Reimann, M., & Schilke, O. (2009). Worldwide faculty perceptions of marketing journals: rankings, trends, comparisons, and segmentations. *Global Edge Business Review, 3*(3), 1-23.

ISACA (2019a). Future of IT Audit Report. Schaumburg, USA.

ISACA (2019b). CISA Review Manual. 27th Edition. Schaumburg, USA.

ISACA, & Protiviti Inc. (2019). 2019 IT Audit Benchmarking Study. Today's toughest challenges in IT audit: Tech Partnerships, Talent, Transformation.

ISACA. (n.d). History [Web page]. Retrieved from https://www.isaca.org/why-isaca/about-us/history

Jans, M., & Hosseinpour, M. (2019). How active learning and process mining can act as continuous auditing catalyst. *International Journal of Accounting Information Systems, 32*, 44-58. https://doi.org/10.1016/j.accinf.2018.11.002

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering - A systematic literature review. *Information and Software Technology, 51*(1), 7–15. https://doi.org/10.1016/j.infsof.2008.09.009

Khan, R., Adi, E., & Hussain, O. (2021). AI-based audit of fuzzy front end innovation using ISO56002. *Managerial Auditing Journal, 36*(4), 564-590. https://doi.org/10.1108/MAJ-03-2020-2588

Krieger, F., Drews, P., & Velte, P. (2021). Explaining the (non-) adoption of advanced data analytics in auditing: A process theory. *International Journal of Accounting Information Systems, 41*, 100511. https://doi.org/10.1016/j.accinf.2021.100511

Kuna, H. D., García-Martinez, R., & Villatoro, F. R. (2014). Outlier detection in audit logs for application systems. *Information Systems, 44*, 22-33. http://dx.doi.org/10.1016/j.is.2014.03.001

Manhanga, L. (2020). *Data analytics in information systems (IS) auditing: An examination of the cost-effectiveness of the use of data analytics in information systems auditing* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 27997995)

Mendez, H. (2020). Innovation in the IT audit process: the IT semantic audit models audit segments using semantic graphs. *ISACA Journal, 1*. Retrieved from https://www.isaca.org/resources/isaca-journal/issues/2020/volume-1/innovation-in-the-it-audit-process

Rozinat, A., & Van der Aalst, W. M. (2008). Conformance checking of processes based on monitoring real behavior. *Information Systems, 33*(1), 64-95. https://doi.org/10.1016/j.is.2007.07.001

Schumann, G., & Gómez, J. M. (2021). Natural language processing in internal auditing - a structured literature review. *AMCIS 2021 Proceedings*. https://aisel.aisnet.org/amcis2021/sig_acctinfosystem_asys/sig_acctinfosystem_asys/1

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Yesmin, T., & Carter, M. W. (2020). Evaluation framework for automatic privacy auditing tools for hospital data breach detections: A case study. *International Journal of Medical Informatics, 138*, 104123. https://doi.org/10.1016/j.ijmedinf.2020.104123

Zerbino, P., Aloini, D., Dulmin, R., & Mininno, V. (2018). Process-mining-enabled audit of information systems: Methodology and an application. *Expert Systems with Applications, 110*, 80-92. https://doi.org/10.1016/j.eswa.2018.05.030

How cite this article

Tanriverdi, N.S., Taskin, N. (2023). A systematic literature review for new technologies in IT audit. *Acta Infologica, 7*(2), 396-408. https://doi.org/10.26650/acin.1142281