

Two decades of financial statement fraud detection literature review; combination of bibliometric analysis and topic modeling approach

Topic
modeling
approach

1367

Milad Soltani and Alexios Kythreotis

Department of Accounting, Economics and Finance, European University Cyprus, Nicosia, Cyprus, and

Arash Roshanpoor

Department of Computer, Yadegar-e-Imam Khomeini (RAH), Janat-abad Branch, Islamic Azad University, Tehran, Iran

Abstract

Purpose – The emergence of machine learning has opened a new way for researchers. It allows them to supplement the traditional manual methods for conducting a literature review and turning it into smart literature. This study aims to present a framework for incorporating machine learning into financial statement fraud (FSF) literature analysis. This framework facilitates the analysis of a large amount of literature to show the trend of the field and identify the most productive authors, journals and potential areas for future research.

Design/methodology/approach – In this study, a framework was introduced that merges bibliometric analysis techniques such as word frequency, co-word analysis and coauthorship analysis with the Latent Dirichlet Allocation topic modeling approach. This framework was used to uncover subtopics from 20 years of financial fraud research articles. Furthermore, the hierarchical clustering method was used on selected subtopics to demonstrate the primary contexts in the literature on FSF.

Findings – This study has contributed to the literature in two ways. First, this study has determined the top journals, articles, countries and keywords based on various bibliometric metrics. Second, using topic modeling and then hierarchy clustering, this study demonstrates the four primary contexts in FSF detection.

Research limitations/implications – In this study, the authors tried to comprehensively view the studies related to financial fraud conducted over two decades. However, this research has limitations that can be an opportunity for future researchers. The first limitation is due to language bias. This study has focused on English language articles, so it is suggested that other researchers consider other languages as well. The second limitation is caused by citation bias. In this study, the authors tried to show the top articles based on the citation criteria. However, judging based on citation alone can be misleading. Therefore, this study suggests that the researchers consider other measures to check the citation quality and assess the studies' precision by applying meta-analysis.

Originality/value – Despite the popularity of bibliometric analysis and topic modeling, there have been limited efforts to use machine learning for literature review. This novel approach of using hierarchical



© Milad Soltani, Alexios Kythreotis and Arash Roshanpoor. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

clustering on topic modeling results enable us to uncover four primary contexts. Furthermore, this method allowed us to show the keywords of each context and highlight significant articles within each context.

Keywords Detection of financial fraud, Latent Dirichlet allocation, Topic modeling, Smart literature review, Bibliometric analysis

Paper type Research paper

1. Introduction

1.1 Description of the research topic

Almost 20 years have passed since the financial crisis caused by Enron's financial fraud. Despite significant reforms such as the Sarbanes Oxley Act and the creation of the Public Company Accounting Oversight Board, which oversees auditors' quality, the news of managers cooking the books continues. Even the legendary investor, the Oracle of Omaha, has fallen victim to financial fraud (Kollewe, 2014). To combat this issue, experts suggest financial insurance (Economist, 2014) and the use of blockchain technology to revolutionize financial transactions (Aicpa.org, 2020). Providing these examples, one may ask, what are the causes of the financial fraud cycle? How can investors protect their capital by examining the quality of their financial statements? Committing financial statement fraud (FSF) is an unethical decision-making process, and detecting fraud is challenging because of its multifaceted nature. For example, 50% of fraud seen by SEC enforcement has been due to revenue recognition (e.g. cookie jar), and 35% is because of concealed liabilities and expenses. The rest is caused by inadequate disclosure in footnotes (Association of Certified Fraud Examiners, 2017).

According to the Association of Certified Fraud Examiners in 2020, three major frauds are corruption, asset misappropriation and FSF. FSF schemes are the least common with the most damaging types of fraud. FSF is the "deliberate misrepresentation of the financial condition of the company accomplished through the intentional misstatement or omission of amount or disclosures of the financial statement" to decisive financial statement users. Therefore, in this study, we intend to review FSF detection to provide a comprehensive picture of the development over 20 years. The applications of the present study can be summarized as follow:

- Uncovering the new trend of articles by analyzing the keywords over two decades;
- Grouping of the authors based upon the similarity of the content of articles;
- Providing the top countries' and journals' collaboration in FSF detection literature; and
- Explaining the primary contexts of FSF literature and providing the high-impact articles in each content based on citation rate.

Therefore, the results of this study will appeal to those interested in comprehensively exploring the FSF detection literature, such as auditors, investors, financial managers and researchers.

2. Literature review

2.1. A review of previous research

In this article, we use a combination of two approaches, bibliometric analysis and topic modeling, to review the research literature on detecting financial fraud. Bibliometric analysis is a quantitative method that we use to analyze a large number of financial fraud studies and uncover emerging trends in FSF detection. This approach is in line with the guideline presented in the article by Donthu *et al.* in 2021, and the software we use is VOSviewer. Table 1 shows the conditions required to use bibliometric analysis.

In the second approach, we will use topic modeling, which aligns with the guideline presented in Asmussen and Møller’s article in 2019. Topic modeling is a natural language processing technique of the unsupervised method that extracts topics from a collection of papers. For topic modeling, we use Python, and the essential Python libraries used in the topic modeling approach include Pemdaz, Gensim and pyLDAvis. [Table 2](#) summarizes the features of the topic modeling approach.

[Table 3](#) shows five studies using bibliometric analysis and topic modeling formatted based on the research field, methodology, required data and data size.

2.2 Research innovation

[Figure 1](#) shows interest over time based on Google Trends data [[1](#)] for the two approaches of bibliometric analysis and topic modeling over the past five years. As is shown, despite the increasing trend for bibliometric analysis, interest in topic modeling has been more significant.

Dataset	Scope	When not to use	When to use	Goal	Methodology
Large	Board	When the high number of papers are heterogeneous or small paper is available	When the scope of review is broad and data is large for manual review	Summarizing a large number of articles and uncovering past and emerging trends	Bibliometric

Table 1. Requirements for using bibliometric analysis

Source: [Donthu et al. \(2021\)](#)

Method	Categories of papers are known?	Coding can be automated?	Person hours spent	Person hours spent interesting results
Topic modeling	No	Yes	Low	Moderate

Table 2. Characteristics of the topic modeling approach

Source: [Asmussen and Møller \(2019\)](#)

Reference	Intended use	Method	Data requirement	Size of data
(Ampese et al., 2022)	Review of anaerobic digestion technology	bibliometric analysis	Academic literature	6,854 articles
(Xu et al., 2021)	Review of entrepreneurship and crisis literature	bibliometric analysis	Academic literature	1,044 articles
(Chen et al., 2021)	Sentic computing	LDA and Bibliometric	Academic literature	308 Articles
(Mustak et al., 2020)	Review of AI research in marketing	LDA and Scientometric analysis	Academic literature	214 Articles
(Sharma et al., 2021)	Information management	LDA and conceptual structure analysis	Academic literature	19,916 Articles

Table 3. Previous studies using bibliometric analysis and topic modeling approaches

Source: Author contribution

We have contributed two ways to the literature on FSF detection. First, we have determined the top journals, articles and keywords based on various bibliometric metrics to show the future trend of the studies. Second, we have demonstrated that the articles with the highest citation rate contribute to detecting FSF literature in different contexts using Latent Dirichlet Allocation (LDA) and clustering methods. To the best of the authors' knowledge, no previous articles in the field of FSF have effectively used bibliometric analysis and topic modeling. Also, by applying hierarchy clustering on subtopics, we demonstrate the main contexts in FSF detection for the first time. Our research questions for the three levels of analysis, i.e. articles, journals and keywords, are as follows:

- RQ1.* What has been the trend in FSF detection article publishing from 2001 to 2021?
- RQ2.* Which countries have the highest number of published articles in the field of FSF detection?
- RQ3.* How do the leading nations collaborate in publishing articles related to FSF detection?
- RQ4.* Which journals have the highest publication rates for articles on FSF detection?
- RQ5.* What has been the trend of the top journals in the FSF detection field from 2001 to 2021?
- RQ6.* Which are the most frequently used keywords in FSF detection articles based on co-occurrence measurement?
- RQ7.* What has been the trend of the top keywords in the field of FSF detection from 2001 to 2021?
- RQ8.* What are the main contexts in FSF detection and which keywords have a higher frequency within these contexts?
- RQ9.* Which articles have the highest citation rate in each context related to FSF detection?

3. Methodology

The methodology of this study is illustrated in [Figure 2](#). Essentially, the research process encompasses three phases. The initial phase, referred to as pre-processing, entails the

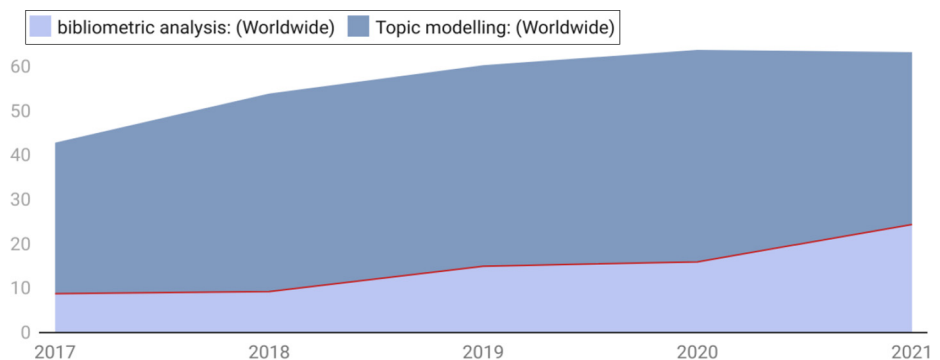


Figure 1.
Trend of interest in bibliometric analysis and topic modelling based on Google Trends

Source: Author contribution

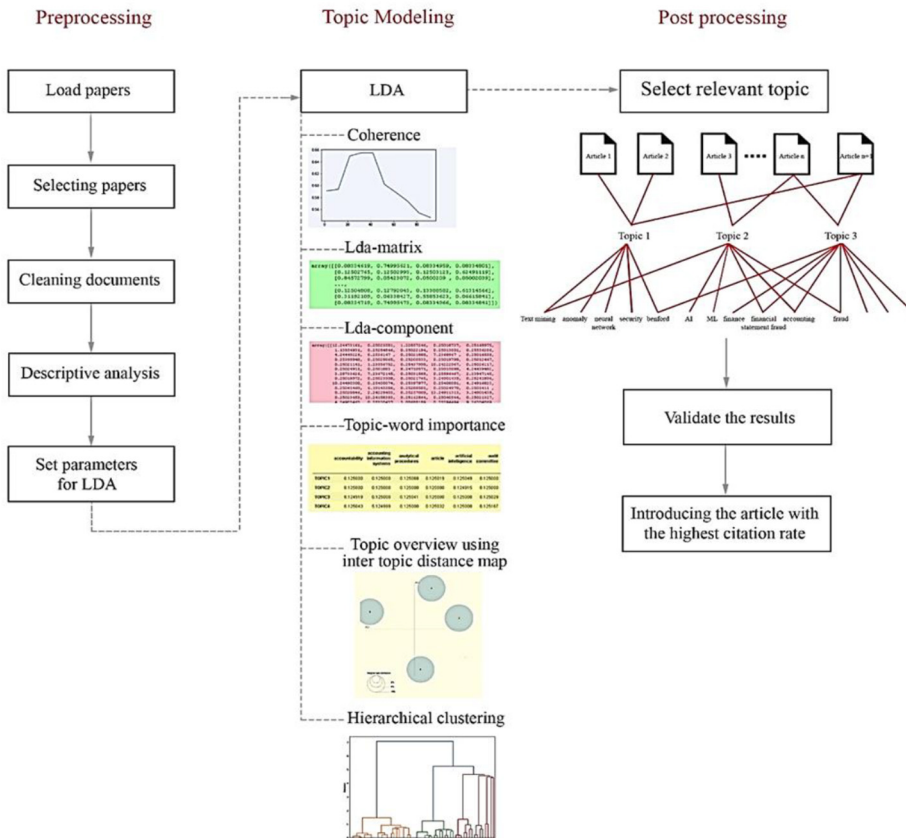


Figure 2. Research methodology to analyze articles in financial statement fraud detection

Source: Author contribution

provision of the relevant data from articles in an Excel format, followed by data cleansing. The next stage is topic modeling, where the LDA method is applied. The clustering method is then used to identify contexts associated with financial fraud detection. Finally, the post-processing stage highlights the most relevant articles for the selected contexts:

- Loading papers: Generally, two primary databases for collecting scholarly publications include Web of Science (WoS) and Scopus. While WoS is one of the largest and most reliable databases for reviewing research literature, the journals in Scopus are more comprehensive (Agarwal, 2016; Saba et al., 2020). Therefore, we have used the Scopus database to find relevant articles.
- Selecting papers: Distinguishing between fraud detection and FSF detection is necessary to collect the relevant articles. To investigate FSF detection, we use the keywords: “anomaly detection financial*” OR “detect fraudulent financial reporting*” OR “detect accounting fraud*” OR “detect misstatements*” OR “financial ratio risk detection*” OR “detect earning manipulation*” for the years 2001 to 2021. The total number of articles is 1,496 articles. Then, 124 articles that are not related to detecting

financial fraud are removed. Also, considering that one of the goals of this study is to categorize research based on author keywords into corresponding contexts, we excluded articles that did not have author keywords. As a result, by removing 296 articles, we selected 1,076 articles suitable for analysis.

- **Descriptive analysis:** In this study, we have used VOSviewer and Phyton software to implement three methods of bibliometric analysis, including co-word analysis, co-bibliometric analysis and co-authorship analysis. Co-word analysis shows that keywords related to each other have the same thematic map. Co-authorship analysis shows the interaction between scholars in detecting financial fraud, such as affiliated institutions and countries. Bibliographic coupling assumes that two publications sharing common references are likely to have a higher similarity.
- **Cleaning documents:** In this step, we break down the keywords into individual units and convert them to lowercase. Additionally, we eliminate any non-alphabetic symbols (such as punctuation marks) and words with less than three characters, as they have minimal impact on the topics being analyzed. Subsequently, we select keywords that do not contain stop words such as “use.” Then, using a process called lemmatization, we identify the root words of the keywords. Finally, we eliminate author keywords that only occur once, as they offer limited value in performing topic modeling.
- **Vectorization:** Following the document cleaning process, we compile a list of keywords indicating the frequency at which each keyword appears in the training set. Subsequently, with the use of the count vectorizer, we convert each keyword into a vector representation.
- **Setting parameters for LDA:** The LDA method has a default value set. In this study, we have changed the parameters of burn-in time, the number of iterations and seed values to achieve better results. In addition, the number of fold parameters has been removed since all the papers are used to run the model.
- **Topic Modeling:** In this study, we used the LDA method, a probabilistic and unsupervised modeling approach. This method allows us to categorize texts within a corpus into specific topics. It is vital at this stage to determine the optimal number of topics through the use of the topic coherence metric. The coherence metric demonstrates how well a topic is supported by the reference corpus (Joao Pedro, 2022). The optimal number of topics for the model can be chosen by visualizing the coherence score, which ranges from 0 to 1 on the vertical axis, with the horizontal axis representing the number of topics. The highest point on the graph indicates the ideal number of topics (Yadav, 2022).
- **LDA model outputs:** The result produced by LDA are the LDA components that demonstrate the significance of each keyword in relation to various topics. The LDA-matrix shows the significance of articles to each topic based on the keyword list used in the research papers. To clarify the outcome, an inter topic distance map has been used which visualizes the value of keywords in different topics based on their frequency of occurrence (Yadav, 2022).
- **Hierarchical clustering:** We use hierarchical clustering to categorize such as topics into a single cluster. Clusters featuring similar topics create a context that is linked to articles. Afterwards, by introducing a new column, we demonstrate the affiliation of each article to its corresponding cluster.

- Selecting the relevant topics: Once the membership of articles within each cluster is established, it becomes imperative to assign labels to each cluster. These labels signify the context of articles that share similar themes. The determination of the labels is based on the occurrence and recognition of significant keywords. Hence, we use the word cloud visualization approach which displays the frequency of keywords in each context.
- Validating the results: The exploratory search ends with labeling the contexts, and it is necessary to check the validity of the results. Validity shows the extent to which the contexts' labeling reflects the collection of keywords in each context. The validity of the results in this study is investigated through semantic validation, in which experts review the keywords in the word cloud to see if the naming of the contexts based on the observed keywords is justified (Asmussen and Møller, 2019). For this purpose, we will ask three finance and financial fraud experts to express their agreement with our labeling. We will also use the following formula to check the overall level of agreement or disagreement (Neuman, 2013):

$$\text{Validity Percentage} = \frac{2 \times \text{number of agreements}}{\text{Total number of codes}} \times 100 \quad (\text{Eqn.1})$$

When the validity percentage in the above formula is higher than 60%, it indicates the proper validity for labeling the contexts:

- Introducing articles with the highest citation rate: In the last step, the top articles in each context are shown based on their citation rates. The citation rate of each article is calculated by dividing the total citations by the number of years that have passed since its publication. This analysis helps us to introduce the most important research topics within the field of study.

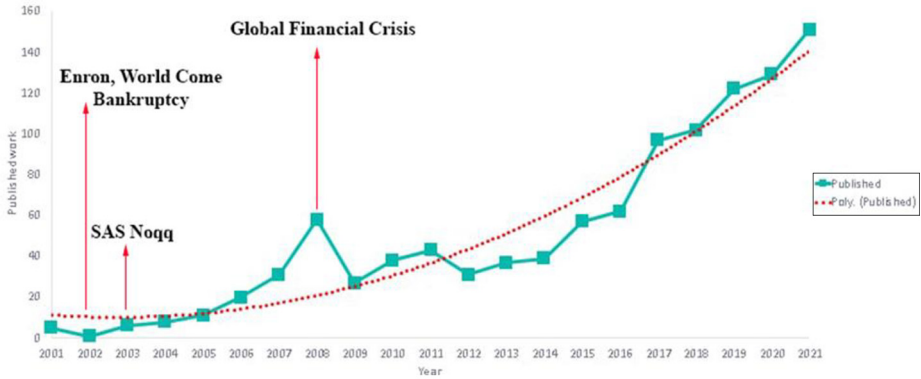
4. Analysis of publications

4.1 Descriptive details of extant publications

4.1.1 *Articles published each year (RQ1)*. Figure 3 shows the frequency of the articles published from 2001 to 2021. The publication trend shows that researchers' interest in issues related to financial fraud detection has been increasing. We have demonstrated this trend with important events related to financial fraud and global markets. For instance, in 2002, Enron and WorldCom went bankrupt due to FSF. One of the consequences of this bankruptcy was the publication of Statement of Auditing Standards No. 99 (SAS No. 99) by the American Institute of Certified Public Accountants. Also, in 2008 the world faced the most significant financial crisis, which put considerable pressure on companies for financing.

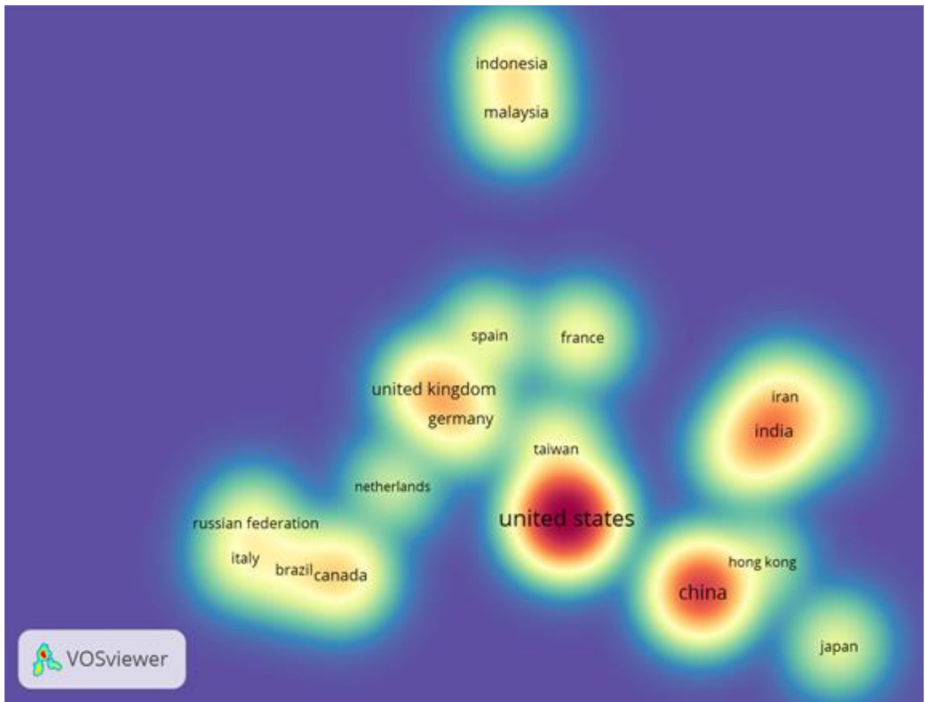
4.1.2 *Density visualization of top 20 countries collaborating in the detection of financial fraud (RQ2)*. Figure 4 illustrates the top 20 countries that have published articles on FSF. The countries included in the data have published at least 14 articles on the topic. The United States of America and China are the leading countries in terms of publishing articles on FSF, with 248 and 134 articles respectively. This suggests that financial fraud is a prevalent issue worldwide and that various countries have contributed to the research in this field.

Figure 3.
Year-wise publication
of detection financial
fraud between 2001
and 2021



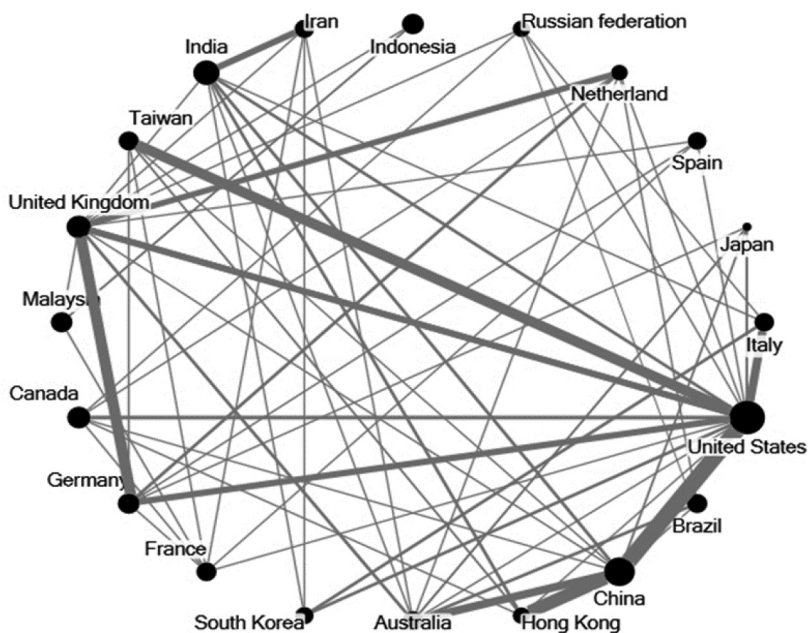
Source: Author contribution

Figure 4.
Top countries
contributing to
publishing financial
statement fraud
detection articles



Source: Author contribution

4.1.3 Network structure of the top 20 countries (RQ3). Figure 5 shows the collaboration between top 20 countries in publishing scholarly papers on FSF detection based on authors' affiliation information. The size of the nodes represents the number of published articles and the width of the edges represents the cooperation between countries. The USA, the UK and



Source: Author contribution

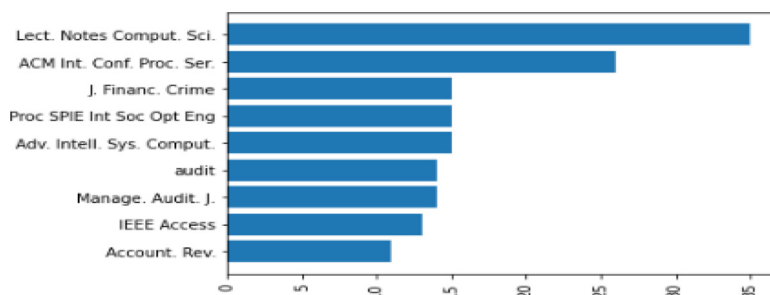
Figure 5. Collaboration map of the top 20 countries in financial statement fraud detection

China have the largest co-authorships while countries with limited cooperation rely on local or individual authors.

4.1.4 Total number of published articles in each journal (RQ4). Figure 6 shows journals that have published more than 10 articles on FSF detection.

As demonstrated, *Lecture Notes in Computer Science* with 35 articles, *ACM International Conference Proceeding Series* with 26 articles and *Journal of Financial Crime* with 15 articles are among the top journals in FSF detection.

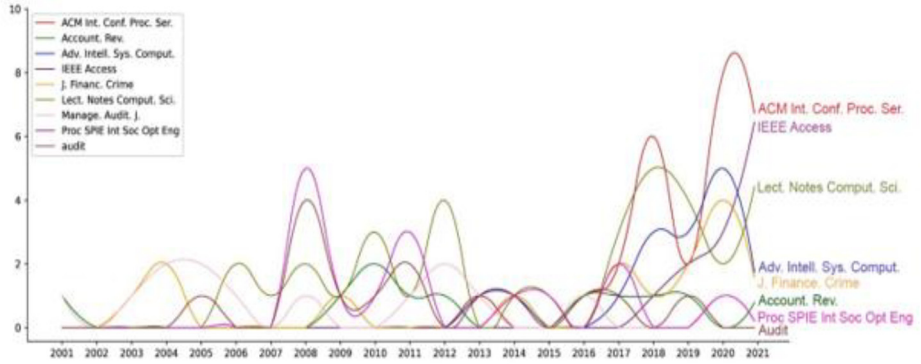
4.1.5 Trend of journals in publishing articles (RQ5). Figure 7 shows the trend of top journals publishing the highest number of articles annually.



Source: Author contribution

Figure 6. Top journals with the highest number of published articles

Figure 7.
The trend of the top journals in the field of financial fraud detection annually



Source: Author contribution

As demonstrated, the journals *Lecture Notes in Computer Science*, *ACM International Conference Proceeding Series*, *IEEE Access* and *Journal of Financial Crime* have an increasing trend from 2001 to 2021.

4.1.6 *Network structure of the top 50 keywords (RQ6).* Figure 8 demonstrates the network structure of the keywords based on the co-occurrence and shows how often each keyword is associated with other keywords. For this purpose, we set the minimum number of co-occurrence to 28, based on which 50 keywords with the highest co-occurrence are shown in Figure 8. Different colors of the nodes represent the initial date of their publication.

The results in Figure 8 show that keywords such as big data, machine learning and decision tree are among the keywords that have recently attracted the attention of

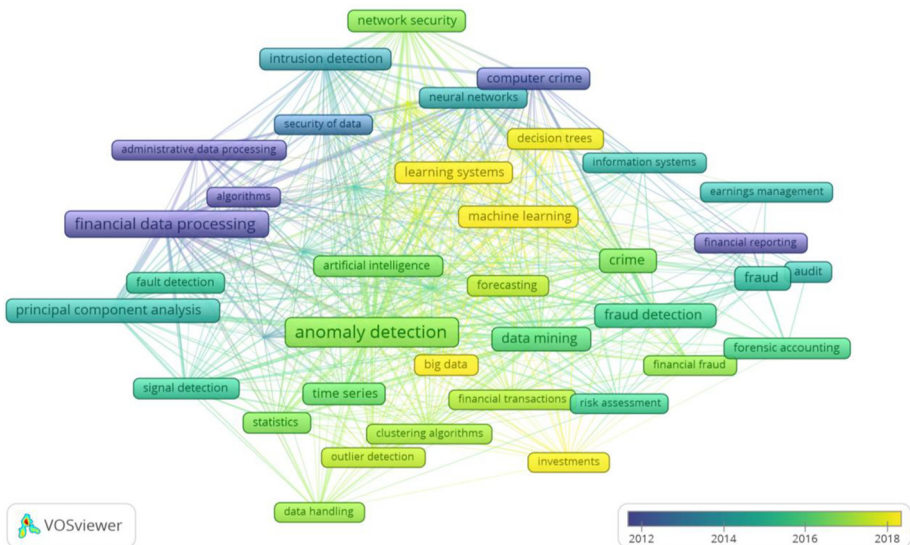


Figure 8.
Network structure for the top 50 keywords concerning financial statement fraud detection

Source: Author contribution

researchers. In contrast, keywords such as computer crime, administrative data processing and financial data processing have been used since 2012.

4.1.7 *Keyword trend (RQ7)*. Table 4 shows the top 20 keywords in FSF detection. It shows that keywords such as anomaly detection, fraud detection, machine learning and data mining have the highest frequency in FSF detection.

In the same line, Figure 9 shows the trend of six keywords based on the frequency of usage from 2001 to 2021. As it can be seen, the keywords of anomaly detection, machine learning and fraud detection have an upward trend, indicating the increasing attention they have received from researchers over the years.

4.2 *Topic modeling using Latent Dirichlet Allocation*

4.2.1 *Coherence scores*. The coherence score is used to determine the optimal number of topics in a reference corpus and was calculated for 100 possible topics. The score reached its

Index	Author's keywords	Frequency
1	Anomaly detection	249
2	Fraud detection	98
3	Machine learning	62
4	Data mining	47
5	Audit	37
6	Principal component analysis	35
7	Earning management	23
8	Deep learning	21
9	Clustering	20
10	Corporate governance	16
11	Unsupervised learning	15
12	Benford's law	15
13	Network security	13
14	Fraud triangle	12
15	Random forest	11
16	Feature extraction	11
17	Neural network	11
18	Audit quality	10
19	Decision tree	10
20	Internal control	10

Table 4. Top keywords based on the frequency criterion in the field of financial fraud detection

Source: Author contribution

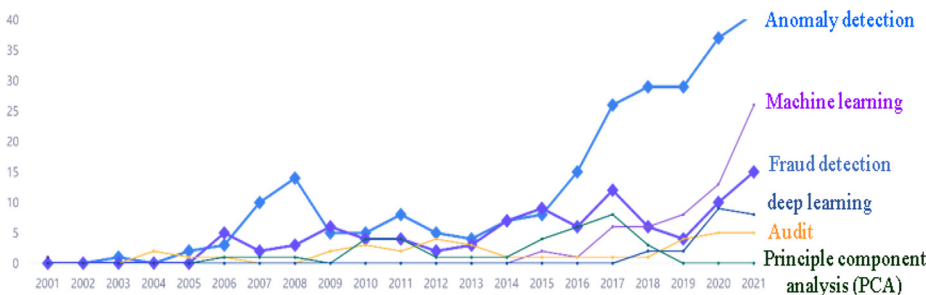


Figure 9. The trend of the number of keywords used over the years 2001–2021

Source: Author contribution

maximum at 0.65, indicating that 42 topics are optimal. Figure 10 represents these 42 topics in a two-dimensional graph, with the intertopic distance map used to evaluate the content of each topic based on its keyword values.

4.2.2 Hierarchical clustering. At this stage, hierarchical clustering is used to reveal the primary contexts related to identifying financial fraud by grouping similar topics. In this method, each article is initially considered a cluster, and at each stage, articles closer to each other form a larger cluster. To find the optimal number of clusters, we have used Dendrogram analysis, a prevalent hierarchical clustering method. The optimal number of clusters is identified in the largest vertical difference between nodes by crossing the horizontal line. Accordingly, the optimal number of clusters is four, and each article belongs to a specific cluster from one to four. Also, four clusters have been formed due to merging similar topics, so it can be concluded that the articles in each of the four clusters refer to different contexts of FSF detection.

Four contexts covering 10%, 80%, 6% and 4.6% of all published articles, respectively. Therefore, contexts two and four have the highest and the lowest number of articles, respectively. In context one, the journals *Lecture Notes in Computer Science*, *Procedia Computer Science* and *ACM International Conference Proceeding Series* have the highest published articles. Moreover, in context two, the journals *Lecture Notes in Computer Science*, *ACM International Conference Proceeding Series*, *Audit* and *Journal of Financial Crime* have the highest number of published articles. In context three, *IEEE Access*, *Lecture Notes in Computer Science*, *Journal of Critical Reviews* and *Studies in Computational Intelligence* have the highest published articles. Finally, in context four, the *Managerial Auditing Journal*, *Issues in Accounting Education* and the *Journal of Financial Crime* have the highest number of published articles. Therefore, *Lecture Notes*

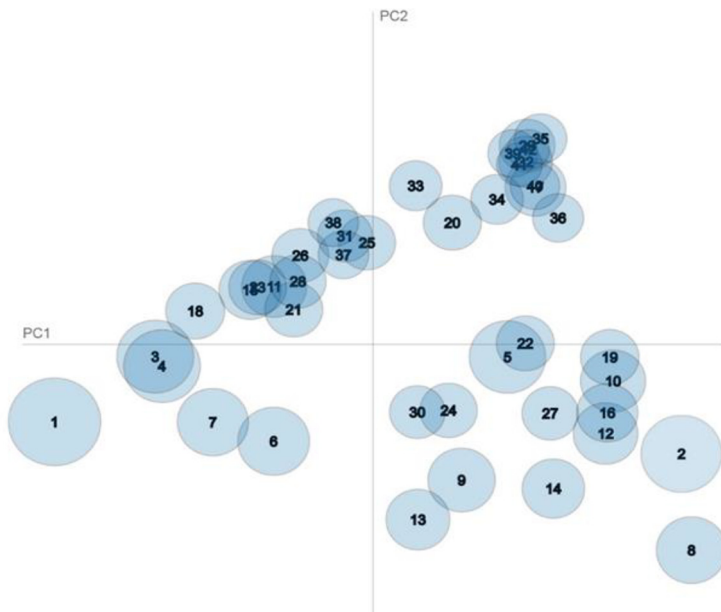


Figure 10.
Intertopic distance
map for the 42
financial statement
fraud detection topics

Source: Author contribution

in *Computer Science* and the *Journal of Financial Crime* have covered various contexts of FSF detection.

Topic modeling approach

4.3 Select relevant topics (RQ8)

The article analyzes the content of various topics based on the frequency of keywords. Out of 3,406 unique keywords, the most repeated keywords in the field of FSF detection from 2001 to 2021 include anomaly detection, fraud, data mining, deep learning audit, clustering audit, Benford law, outlier detection and machine learning. These keywords are shown in the word cloud graph in Figure 11.

4.3.1 Selected relevant labels for context one. In this part, we aim to select relevant labels based on the frequency of keywords in each context. Figure 12 shows the keywords leading to the creation of context one.

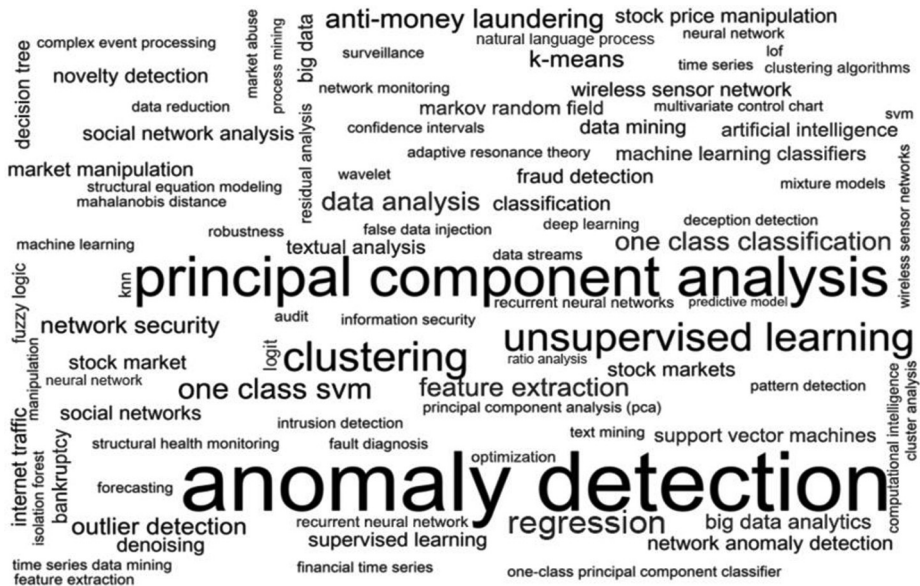
Two FSF detection methods are human detection and machine detection. The results in Figure 13 show that by applying topic modeling, the keywords related to machine detection are more similar and, therefore, are in context one. In other words, human detection, such as the Whistleblowing system, can be a useful tool in detecting FSF, as it can help to bring attention to fraudulent activity that might otherwise go undetected. However, human detection is not the only method of financial fraud detection, and it is not always a reliable or effective method on its own. Financial fraud detection often involves combining different techniques, such as data analysis and machine learning algorithms.

The first group of keywords consists of feature selection, principal component analysis, feature extraction and dimensionality reduction. Generally, the issues related to FSF detection include the study of extensive financial data where the identification of financial variables and financial ratios is necessary. Then, by applying data mining techniques, organizations are classified into two categories fraudulent and non-fraudulent. However, if



Figure 11. Visualization of 3,406 unique keywords in the field of financial statement fraud

Source: Author contribution



Source: Author contribution

Figure 12. Selecting the title of “fraud detection techniques” for context one

the data set includes many irrelevant and correlated features, a curse of dimensionality will appear, reducing the classification’s performance. Therefore, removing the number of irrelevant features from the data set is necessary by using dimensionality reduction techniques such as feature selection and principal component analysis (see, for example, [Gupta and Mehta, 2020](#)).

The second group of keywords consists of text mining, neural network, classification, one class classification, artificial intelligence, time series analysis, graph mining, visual analytics, random forest, regression, unsupervised learning, decision tree, *k*-means, fuzzy logic, supervised learning, time series prediction and correlation. Generally, there are two supervised and unsupervised learning approaches for artificial intelligence and machine learning. In supervised learning, data sets have labels and include classification algorithms (e.g. support vector machine, decision tree and random forest) and regression algorithms (e.g. linear regression and logistic regression). The unsupervised learning approach analyzes unlabeled data sets and includes methods such as clustering and association (see, for example, [Ashtiani and Raahemi, 2021](#)). According to these explanations, it can be argued that the keywords of the first context are related to the title of “fraud detection techniques” for cluster one.

4.3.2 Selected relevant label for context two. [Figure 13](#) shows the frequency of keywords leading to the creation of context two.

Cluster two includes keywords such as fraudulent financial reporting, FSF, earning management, corporate governance, cooking the books, fraud prevention, fraud triangle, fraud diamond, the pentagon model. The bankruptcy of companies such as Enron and WorldCom increased attention to the quality of financial reports, and researchers began to investigate the causal factors associated with the increased probability of fraud and the consequences of financial fraud (see, for example, [Rezaee, Z., 2005](#)). Other studies have tried



Figure 13. Selecting the title of “causes and deterrence of financial statement fraud” for context two

Source: Author contribution

to investigate different proxies to measure fraud factors in the fraud triangle, diamond and pentagon model (see, for example, Skousen et al., 2009). Based on the provided explanations, we select the title “Causes and deterrence of FSF” for Cluster 2.

4.3.3 Selected relevant label for context three. Figure 14 shows the keywords leading to the creation of context three based on the frequency of keywords.

Cluster three includes the keywords of digital forensics, network security, wireless sensor network, cloud computing, data privacy, malware, DDoS, information security, cyber security, online transaction, website defacement attack. These keywords are related to transactions and frauds caused by computers and the internet, the disclosure of which leads to the loss of the intellectual property of the company secrets and major financial damage. Articles in this area can include a variety of topics related to digital forensics. Cyber security measures can be an important tool in preventing and detecting fraud in financial statements. For example, implementing strong access controls, monitoring financial transactions and maintaining accurate audit logs can help identify manipulation of financial data. Also, digital forensics can be used to analyze the digital evidence that has been collected, to find and recover hidden information. For example, studies focuses on data manipulation detection methods, which means the unauthorized modification of the system (e.g. data leakage malware, salami technique) to disrupt the normal function of the targeted system (see, for example, Nicholls et al., 2021). Other study objectives can be identifying factors related to preventing online fraud and data security, such as security auditing and data classification (see, for example, Soomro et al., 2021). Based on the provided explanations, we select the title “computer and online transaction fraud” for cluster three.

external audits. Internal auditors are better positioned to detect financial fraud due to their proximity and understanding (Association of Certified Fraud Examiners, 2017). Therefore, the first group of studies has addressed the role of internal auditors in risk management (see, for example, Tamimi, 2021). In this regard, another group of studies has examined the function of internal auditors under the influence of mediating factors such as corporate governance and gender diversity (see, for example, Pazarskis et al., 2021). The second group of studies is related to external auditors. For instance, some studies have investigated the effect of external auditor quality (e.g. the auditor’s quality) on the possibility of identifying financial fraud (see, for example, Qawqzeh et al., 2021). Other studies have investigated new trends in auditing financial statements (see, for example, Lim, 2021).

4.4 Validity test of the labeling for the contexts

At this stage, it is necessary to examine the validity of labeling for the four mentioned contexts. To this end, we provided the keywords of each topic to the financial experts and asked them to express their agreement with the keywords related to each context. If the calculated value of validating percentage is more than 60%, the created context will have sufficient validity (Neuman, 2013). Table 5 shows the level of agreement of labeling based on the keywords of each context by three financial and financial fraud experts.

4.5 Introducing top articles for main contexts in financial statement fraud detection (RQ9)

In this section, the top articles of each context are introduced in Table 6. For this purpose, the citation rate index is used, and the articles are sorted based on this index. Articles with context number one are related to fraud detection techniques. Articles with context number two are related to the causes and deterrence of FSF. Articles with context number three are related to computer and online transaction fraud. Articles with context number four are related to auditors’ fraud-related responsibilities.

5. Conclusion

This study aims to review the research literature concerning FSF detection from 2001 to 2021. Accordingly, 1,496 articles were extracted from the Scopus database. After

Interview no.	Title context	Context no.	The total no. of keywords in each cluster	No. of approvals keywords	Validity of labeling (%)
1	Fraud detection techniques	1	244	90	73
2				86	70
3				86	70
1	Causes and deterrence of financial statement fraud	2	2,834	885	62
2				865	61
3				870	61
1	Computer and online transaction fraud	3	170	70	82
2				65	76
3				70	82
1	Auditors’ fraud-related responsibilities	4	158	75	94
2				68	86
3				60	75

Table 5. Validity test of the labeling for the contexts

Source: Author contribution

Context no.	Citation	Citation rate	Title	Author	Year
2	678	96.85	Graph based anomaly detection and description: A survey	Akoglu L., <i>et al.</i>	2015
1	166	55.33	Real-time big data processing for anomaly detection: A Survey	Habeeb R.A., <i>et al.</i>	2019
2	32	32	An Integrated Cluster Detection, Optimization, and Interpretation Approach for Financial Data	Li T., Kou G., Peng Y., Yu P.S.	2021
2	171	28.5	Intelligent financial fraud detection: A comprehensive review	West J., Bhattacharya M.	2016
1	45	22.5	Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach	Bao Y., <i>et al.</i>	2020
2	154	15.4	The evolution of fraud theory	Dorminey J., <i>et al.</i>	2012
4	36	12	The role of audit in the fight against corruption	Jeppesen	2019
4	67	9.57	Materiality guidance of the major public accounting firms	Eilifsen A., Messier W.F., Jr.	2015
4	99	9	Financial statement fraud detection: An analysis of statistical and machine learning algorithms	Perols J.	2011
1	169	8.89	Earnings Manipulation in Failing Firms	Rosner R.L.	2003
1	53	8.83	Unsupervised learning for robust Bitcoin fraud detection	Monamo P., <i>et al.</i>	2016
3	26	8.66	Situ: Identifying and explaining suspicious behavior in networks	Goodall J.R., <i>et al.</i>	2019
3	33	8.25	Malware analysis and detection using data mining and machine learning classification	Chowdhury M., <i>et al.</i>	2018
4	96	8	The world has changed - Have analytical procedure practices?	Trompeter G., Wright A.	2010
3	17	5	A flow-based approach for Trickbot banking trojan detection	Gezer A., <i>et al.</i>	2019
3	12	4	Stock Price Manipulation Detection using Generative Adversarial Networks	Leangarun T., <i>et al.</i>	2019

Table 6.
Top articles for main contexts in financial fraud detection

Source: Author contribution

screening the articles, we selected 1,076 papers for analysis. To analyze the literature on FSF detection, we first used the bibliometric approach and revealed that the articles published during the past two decades have an upward trend. We also indicated the top 20 countries with the highest number of articles published. The results showed that the USA and China are the leading countries in content production in FSF detection. Also, the USA and China coauthors cooperate more with other countries. Then, we reviewed the journals and identified the top journals with more than 10 articles during the past two decades. Our results also revealed that the journals of *Lecture Notes in Computer Science*, *ACM International Conference Proceeding Series*, *IEEE Access* and *Journal of Financial Crime* have an increasing trend in publishing content FSF detection. Then, we analyzed the keywords and showed that keywords such as decision trees, machine learning and big data have recently attracted researchers' attention. Then, we introduced the 20 most frequent keywords in the literature on FSF detection. The analysis of the trends of the keywords such as anomaly detection, machine learning and fraud detection showed a growing trend in using these keywords in recent years.

Finally, using the LDA modeling method, 42 related topics were identified. Finally, we identified the contexts by applying hierarchical clustering. The examination of four clusters revealed that the *Journal of Financial Crime* and *Lecture Notes in Computer Science* include more diverse topics because of their presence in most contexts. Then, using word cloud, we displayed the keywords of each context and identified the four labels of fraud detection techniques, fraud prevention and deterrence, computer and online transaction fraud and auditors' fraud-related responsibilities based on the analysis. Finally, we have introduced the top articles in each context label based on citation rate.

6. Limitation of the study

In this study, we tried to comprehensively view the studies related to financial fraud conducted over two decades. However, this research has limitations that can be an opportunity for future researchers. The first limitation is due to language bias. We have focused on English language articles, so it is suggested that other researchers consider other languages as well. The second limitation is caused by citation bias. In this study, we tried to show the top articles based on the citation criteria. However, judging based on citation alone can be misleading. Therefore, we suggest that the researchers consider other measures to check the citation quality and assess the studies' precision by applying meta-analysis.

Note

1. <https://trends.google.com/trends/>

References

- Agarwal, A., Durairajanayagam, D., Tatagari, S., Esteves, S., Harlev, A., Henkel, R., Roychoudhury, S., Homa, S., Puchalt, N., Ramasamy, R., Majzoub, A., Ly, K., Tvrda, E., Assidi, M., Kesari, K., Sharma, R., Banihani, S., Ko, E., Abu-Elmagd, M. and Gosalvez, J. (2016), "Bibliometrics: tracking research impact by selecting the appropriate metrics", *Asian Journal of Andrology*, Vol. 18 No. 2, p. 296, (online), doi: [10.4103/1008-682x.171582](https://doi.org/10.4103/1008-682x.171582).
- Aicpa.org (2020), "Blockchain versus financial statement fraud", (online), available at: www.aicpa.org/professional-insights/download/blockchain-versus-financial-statement-fraud (accessed 17 August 2022).

- Ampese, L.C., Sganzerla, W.G., Di Domenico Ziero, H., Mudhoo, A., Martins, G. and Forster-Carneiro, T. (2022), "Research progress, trends, and updates on anaerobic digestion technology: a bibliometric analysis", *Journal of Cleaner Production*, Vol. 331, p. 130004, (online), doi: [10.1016/j.jclepro.2021.130004](https://doi.org/10.1016/j.jclepro.2021.130004).
- Ashtiani, M.N. and Raaheemi, B. (2021), "Intelligent fraud detection in financial statements using machine learning and data mining: a systematic literature review", *IEEE Access*, Vol. 10, p. 1, (online), doi: [10.1109/access.2021.3096799](https://doi.org/10.1109/access.2021.3096799).
- Asmussen, C.B. and Møller, C. (2019), "Smart literature review: a practical topic modeling approach to exploratory literature review", *Journal of Big Data*, Vol. 6 No. 1, pp. 1-18, (online), doi: [10.1186/s40537-019-0255-7](https://doi.org/10.1186/s40537-019-0255-7).
- Association of Certified Fraud Examiners (2017), *Fraud Examiner's Manual*, Association of Certified Fraud Examiners, Austin, TX.
- Chen, X., Xie, H., Cheng, G. and Li, Z. (2021), "A decade of sentic computing: topic modeling and bibliometric analysis", *Cognitive Computation*, Vol. 14 No. 1, pp. 24-47, (online), doi: [10.1007/s12559-021-09861-6](https://doi.org/10.1007/s12559-021-09861-6).
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. and Lim, W.M. (2021), "How to conduct a bibliometric analysis: an overview and guidelines", *Journal of Business Research*, Vol. 133, pp. 285-296, (online), doi: [10.1016/j.jbusres.2021.04.070](https://doi.org/10.1016/j.jbusres.2021.04.070).
- Gupta, S. and Mehta, S.K. (2020), "Feature selection for dimension reduction of financial data for detection of financial statement frauds context to Indian companies", *Global Business Review*, (online), doi: [10.1177/0972150920928663](https://doi.org/10.1177/0972150920928663).
- Kollewe, J. (2014), "Warren buffett: 'tesco was a huge mistake'", *The Guardian*, Retrieved from www.theguardian.com/business/2014/oct/02/warren-buffet-tesco-huge-mistake
- Lim, F.X. (2021), "Emerging technologies to detect fraud in audit testing: a perception of Malaysian big four auditors", *SSRN Electronic Journal*, (online), doi: [10.2139/ssrn.3877347](https://doi.org/10.2139/ssrn.3877347).
- Mustak, M., Salminen, J., Plé, L. and Wirtz, J. (2020), "Artificial intelligence in marketing: topic modeling, scientometric analysis, and research agenda", *Journal of Business Research*, Vol. 124, pp. 389-404, (online), doi: [10.1016/j.jbusres.2020.10.044](https://doi.org/10.1016/j.jbusres.2020.10.044).
- Neuman, W.L. (2013), *Workbook for Social Research Methods: qualitative and Quantitative Approaches*, 7th ed., Allyn And Bacon, Pearson Higher Education, Boston.
- Nicholls, J., Kuppa, A. and Le-Khac, N.A. (2021), "Financial cybercrime: a comprehensive survey of deep learning approaches to tackle the evolving financial crime landscape", *IEEE Access*, Vol. 9, pp. 163965-163986, (online), doi: [10.1109/access.2021.3134076](https://doi.org/10.1109/access.2021.3134076).
- Pazarskis, M., Koutoupis, A.G., Kyriakou, M.I. and Galanis, S. (2021), "Corporate governance and internal audit at Greek municipal enterprises in the COVID-19 era", in *Proceedings of the Corporate Governance: An Interdisciplinary Outlook in the Wake of Pandemic Conference, 19-20 November 2020*, Virtus interpress, Sumy, Ukraine, (online), pp. 142-146, doi: [10.22495/cgsetpt21](https://doi.org/10.22495/cgsetpt21).
- Pedro, J. (2022), "[review of understanding topic coherence measures] from towards data science", (online), available at: <https://towardsdatascience.com/understanding-topic-coherence-measures-4aa41339634c> (accessed 19 August 2022).
- Qawqzeh, H.K., Endut, W.A. and Rashid, N. (2021), "Board components and quality of financial reporting: mediating effect of audit quality", *Journal of Contemporary Issues in Business and Government*, Vol. 27 No. 2, p. 179, (online), doi: [10.47750/cibg.2021.27.02.023](https://doi.org/10.47750/cibg.2021.27.02.023).
- Rezaee, Z. (2005), "Causes, consequences, and deterrence of financial statement fraud", *Critical Perspectives on Accounting*, Vol. 16 No. 3, pp. 277-298, doi: [10.1016/s1045-2354\(03\)00072-8](https://doi.org/10.1016/s1045-2354(03)00072-8).
- Saba Ebrahim, Poshtan, J., Jamali, S.M. and Ebrahim, N.A. (2020), "Quantitative and qualitative analysis of time-series classification using deep learning", *IEEE Access*, Vol. 8, pp. 90202-90215.
- Sharma, A., Rana, N.P. and Nunkoo, R. (2021), "Fifty years of information management research: a conceptual structure analysis using structural topic modeling", *International Journal of Information Management*, Vol. 58, p. 102316, (online), doi: [10.1016/j.ijinfomgt.2021.102316](https://doi.org/10.1016/j.ijinfomgt.2021.102316).

- Skousen, C.J., Smith, K.R. and Wright, C.J. (2009), "Detecting and predicting financial statement fraud: the effectiveness of the fraud triangle and SAS no. 99", in *Corporate Governance and Firm Performance*, Emerald Group Publishing Limited, Bingley, USA, Vol. 13, pp. 53-81.
- Soomro, Z.A., Shah, M.H. and Thatcher, J. (2021), "A framework for ID fraud prevention policies in the E-tailing sector", *Computers and Security*, Vol. 109, p. 102403, (online), doi: [10.1016/j.cose.2021.102403](https://doi.org/10.1016/j.cose.2021.102403).
- Tamimi, O. (2021), "The role of internal audit in risk management from the perspective of risk managers in the banking sector", *Australasian Business, Accounting and Finance Journal*, Vol. 15 No. 2, pp. 114-129, (online), doi: [10.14453/aabfj.v15i2.8](https://doi.org/10.14453/aabfj.v15i2.8).
- The Economist (2014), "The dozy watchdogs", Retrieved from www.economist.com/briefing/2014/12/11/the-dozy-watchdogs
- Xu, Z., Wang, X., Wang, X. and Skare, M. (2021), "A comprehensive bibliometric analysis of entrepreneurship and crisis literature published from 1984 to 2020", *Journal of Business Research*, Vol. 135, pp. 304-318, (online), doi: [10.1016/j.jbusres.2021.06.051](https://doi.org/10.1016/j.jbusres.2021.06.051).
- Yadav, K. (2022), "The complete practical guide to topic modelling", (online) Medium. (online), available at: <https://towardsdatascience.com/topic-modelling-f51e5ebfb40a>

Further reading

- Akoglu, L., Tong, H. and Koutra, D. (2015), "Graph-based anomaly detection and description: a survey", *Data Mining and Knowledge Discovery*, Vol. 29 No. 3, pp. 626-688, (online), doi: [10.1007/s10618-014-0365-y](https://doi.org/10.1007/s10618-014-0365-y).
- Bao, Y., Ke, B., Li, B., Yu, Y.J. and Zhang, J. (2020), "Detecting accounting fraud in publicly traded US firms using a machine learning approach", *Journal of Accounting Research*, Vol. 58 No. 1, pp. 199-235, (online), doi: [10.1111/1475-679X.12292](https://doi.org/10.1111/1475-679X.12292).
- Chowdhury, M., Rahman, A. and Islam, R. (2017), "Malware analysis and detection using data mining and machine learning classification", *International conference on applications and techniques in cyber security and intelligence*, (online), pp. 266-274 doi: [10.1007/978-3-319-67071-3_33](https://doi.org/10.1007/978-3-319-67071-3_33).
- Dorminey, J., Fleming, A.S., Kranacher, M.J. and Riley, R.A. Jr. (2012), "The evolution of fraud theory", *Issues in Accounting Education*, Vol. 27 No. 2, pp. 555-579, (online), doi: [10.2308/iaee-50131](https://doi.org/10.2308/iaee-50131).
- Ebrahim, S.A., Poshtan, J., Jamali, S.M. and Ebrahim, N.A. (2020), "Quantitative and qualitative analysis of time-series classification using deep learning", *IEEE Access*, Vol. 8, pp. 90202-90215, (online), doi: [10.1109/access.2020.2993538](https://doi.org/10.1109/access.2020.2993538).
- Eilifsen, A. and Messier, W.F. Jr (2015), "Materiality guidance of the major public accounting firms", *Auditing: A Journal of Practice and Theory*, Vol. 34 No. 2, pp. 3-26, (online), doi: [10.2308/ajpt-50882](https://doi.org/10.2308/ajpt-50882).
- Gezer, A., Warner, G., Wilson, C. and Shrestha, P. (2019), "A flow-based approach for Trickbot banking Trojan detection", *Computers and Security*, Vol. 84, pp. 179-192, (online), doi: [10.1016/j.cose.2019.03.013](https://doi.org/10.1016/j.cose.2019.03.013).
- Goodall, J.R., Ragan, E.D., Steed, C.A., Reed, J.W., Richardson, G.D., Huffer, K.M., Bridges, R.A. and Laska, J.A. (2018), "Situ: identifying and explaining suspicious behavior in networks", *IEEE Transactions on Visualization and Computer Graphics*, Vol. 25 No. 1, pp. 204-214, (online), doi: [10.1109/TVCG.2018.2865029](https://doi.org/10.1109/TVCG.2018.2865029).
- Habeeb, R.A.A., Nasaruddin, F., Gani, A., Hashem, I.A.T., Ahmed, E. and Imran, M. (2019), "Real-time big data processing for anomaly detection: a survey", *International Journal of Information Management*, Vol. 45, pp. 289-307.
- Jeppesen, K.K. (2019), "The role of auditing in the fight against corruption", *The British Accounting Review*, Vol. 51 No. 5, p. 100798, (online), doi: [10.1016/j.bar.2018.06.001](https://doi.org/10.1016/j.bar.2018.06.001).
- Leangarun, T., Tangamchit, P. and Thajchayapong, S. (2018), "Stock price manipulation detection using generative adversarial networks", *2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE, (online), pp. 2104-2111, doi: [10.1109/SSCI.2018.8628777](https://doi.org/10.1109/SSCI.2018.8628777).

- Li, T., Kou, G., Peng, Y. and Philip, S.Y. (2021), "An integrated cluster detection, optimization, and interpretation approach for financial data", *IEEE Transactions on Cybernetics*, Vol. 52 No. 12, pp. 13848-13861, (online), doi: [10.1109/TCYB.2021.3109066](https://doi.org/10.1109/TCYB.2021.3109066).
- Madahali, L. and Hall, M. (2020), "Application of the Benford's law to social bots and information operations activities", *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, *IEEE*, (online), pp. 1-8, doi: [10.1109/CyberSA49311.2020.9139709](https://doi.org/10.1109/CyberSA49311.2020.9139709).
- Monamo, P., Marivate, V. and Twala, B. (2016), "Unsupervised learning for robust bitcoin fraud detection", *2016 Information Security for South Africa (ISSA)*, *IEEE*, (online), pp. 129-134, doi: [10.1109/ISSA.2016.7802939](https://doi.org/10.1109/ISSA.2016.7802939).
- Perols, J. (2011), "Financial statement fraud detection: an analysis of statistical and machine learning algorithms", *AUDITING: A Journal of Practice and Theory*, Vol. 30 No. 2, pp. 19-50, (online), doi: [10.2308/ajpt-50009](https://doi.org/10.2308/ajpt-50009).
- Rosner, R.L. (2003), "Earnings manipulation in failing firms", *Contemporary Accounting Research*, Vol. 20 No. 2, pp. 361-408.
- The dozy (2014), "Watchdogs", *The Economist*, (online), available at: www.economist.com/briefing/2014/12/11/the-dozy-watchdogs
- Trompeter, G. and Wright, A. (2010), "The world has changed—have analytical procedure practices?", *Contemporary Accounting Research*, Vol. 27 No. 2, pp. 669-700, (online), doi: [10.1111/j.1911-3846.2010.01023_8.x](https://doi.org/10.1111/j.1911-3846.2010.01023_8.x).
- Uuganbayar, G., Yautsiukhin, A., Martinelli, F. and Massacci, F. (2021), "Optimization of cyber insurance coverage with the selection of cost-effective security controls", *Computers and Security*, Vol. 101, p. 102121, (online), doi: [10.1016/j.cose.2020.102121](https://doi.org/10.1016/j.cose.2020.102121).
- West, J. and Bhattacharya, M. (2016), "Intelligent financial fraud detection: a comprehensive review", *Computers and Security*, Vol. 57, pp. 47-66, (online), doi: [10.1016/j.cose.2015.09.005](https://doi.org/10.1016/j.cose.2015.09.005).
- Zhou, C. and Paffenroth, R.C. (2017), "Anomaly detection with robust deep autoencoders", *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (online), doi: [10.1145/3097983.3098052](https://doi.org/10.1145/3097983.3098052).

Corresponding author

Milad Soltani can be contacted at: soltani.milad@gmail.com