OmniaScience

Intangible Capital

IC, 2026 – 22(1): 101-122 – Online ISSN: 1697-9818 – Print ISSN: 2014-3214

https://doi.org/10.3926/ic.3119

Bibliometric analysis of big data applications in accounting fraud detection

Carme Monserrat

Barcelona University (Spain)

carmemonserrat@ub.edu

Received October, 2024

Accepted January, 2026

Abstract

Purpose: The use of Big Data can be strategic in the prevention and detection of accounting fraud, helping to protect the integrity of accounting and improving the efficiency of processes. Using these advanced technologies, organisations can minimize their exposures to fraud and maintain more accurate and transparent accounting.

Design/methodology/approach: A bibliometric study was conducted, using a quantitative method based on indicators and on an analysis of the links and connections and visualisation of the network to see the impact, citation, institutions, authors, countries, etc., as well as collaboration between the various agents.

Findings: Research on Big Data in fraud detection has grown rapidly, integrating Machine Learning, Data Mining, and Artificial Intelligence into auditing. The field is global and interdisciplinary, shifting from traditional anomaly checks to predictive, real-time prevention while acknowledging emerging digital fraud risks.

Research limitations/implications: The selection from the Scopus database was made using 6 keywords; the results could have been different if a different or larger set of keywords had been applied to the research carried out.

Practical implications: The findings provide valuable insights into the academic landscape surrounding this topic. Advanced techniques such as Data Mining, Machine Learning, and Deep Learning can significantly enhance transaction transparency and traceability, thereby bolstering confidence in accounting practices and mitigating fraud risks.

Social implications: Fraud is inherently unethical and presents significant challenges for finance and society; technology plays an important role in its detection and prevention. However, reinforcing ethical and sustainability principles is increasingly important. This issue transcends mere technological application; it requires fostering a culture of integrity, transparency, and responsibility.

Originality/value: Through an integrated bibliometric approach, this study provides original value by mapping how Big Data is reshaping fraud-detection research and by highlighting the persistent gap between advanced analytics and audit-standard alignment, underscoring the need for explainable and audit-ready tools.

Keywords: Fraud, Accounting, Big data, Machine learning, Data mining, Forensic accounting

Jel Codes: M49

To cite this article:

Monserrat, C. (2026). Bibliometric analysis of big data applications in accounting fraud detection. *Intangible Capital,* 22(1), 101-122. https://doi.org/10.3926/ic.3119

1. Introduction

In today's digital era, the exponential growth of data generation and computing power (Sandhu, 2022) offers unprecedented opportunities for analysis, including the detection of accounting fraud. Big Data enables auditors and regulators to identify hidden patterns and anomalies through techniques such as Data Mining and Machine Learning (Bao, Hilary & Ke, 2022). Accounting fraud, rooted in deceptive manipulation of financial information, has wide-ranging economic and social consequences (Amat-Salas & Lloret-Gual, 2020). The expansion of global markets and complex financial instruments has amplified these risks, while the post-pandemic digitization of business has further increased exposure to fraud. Against this backdrop, Big Data emerges as a key ally for ensuring the integrity of financial reporting and supporting Forensic Accounting investigations.

Fraud, whose etymological root derives from the Latin *fraus* (gen. *fraudis*), represents an act contrary to truth and honesty that harms those it targets. Within accounting, it manifests as a deliberate manipulation of financial information that undermines trust and distorts the transparency of markets. Beyond its technical dimension, accounting fraud has deep ethical implications, threatening the reliability of organisations and the stability of the economy. Its impact is felt across the financial ecosystem—by investors, banks, suppliers, and public authorities alike (Amat-Salas & Lloret-Gual, 2020).

In this complex scenario, *Big Data* has emerged as a transformative ally. By enabling the analysis of vast volumes of information, it allows hidden patterns and anomalies to surface—clues that may reveal fraudulent behaviour and trigger early preventive action (Saleh, Marei, Ayoush & Abu-Afifa, 2023). The need for such tools has grown sharply: the expansion of global markets, the increasing sophistication of financial instruments, and the speed of digital transactions have all multiplied the opportunities for manipulation. Moreover, the pandemic and the widespread adoption of remote working have accelerated digitalization, creating new vulnerabilities and redefining how fraud operates (Amat-Salas, 2017; Gay-de-Liébana-Saludas, 2013).

At the heart of this technological revolution lies the essence of *Big Data*, defined by its three "Vs": volume, variety, and velocity. These dimensions allow for the processing of enormous quantities of structured and unstructured data, generated at high speed and from multiple sources (Yang & Zhang, 2023). Through this continuous stream of information, organisations can observe financial behaviour in real time, detecting irregularities that traditional methods might overlook.

This analytical capacity is amplified by *Data Mining*, which delves into large data sets to uncover trends and relationships that support strategic decision-making and predictive analysis (Selvarajan, 2021). When combined with *Forensic Accounting*—the discipline dedicated to collecting and presenting financial, legal, and administrative evidence—these tools transform how fraud is investigated (Kayed & Al-Sartawi, 2024). What was once a reactive process focused on proving misconduct after the fact is becoming a proactive, technology-driven system capable of anticipating it.

Regulation is evolving in parallel. The International Auditing Standards adapted for Spain [NIA-ES] (Instituto de Contabilidad y Auditoría de Cuentas–ICAC, 2024) establish that auditors must obtain reasonable assurance that financial statements are free from material misstatements, whether due to error or fraud (NIA-ES 200 and NIA-ES 240). Under traditional approaches, as outlined in NIA-ES 530 "Audit sampling," auditors relied on representative samples to draw conclusions about the whole population. Today, *Artificial Intelligence* and *Big Data* technologies make it possible to analyse the entire universe of transactions, increasing precision and reducing uncertainty (Chen, Wu & Yan, 2022).

This convergence of technology, regulation, and professional judgment defines a new paradigm in fraud detection. *Big Data* not only strengthens auditors' ability to uncover irregularities but also enhances efficiency and

transparency across financial reporting. Advanced *Data Mining* and *Artificial Intelligence* techniques provide real-time alerts, transforming the auditor's role from verifier to continuous risk monitor (Meiryani, Andini, Fahlevi, Yadiati, Purnomo & Prajena, 2022).

Within this context, the present paper conducts a bibliometric study to trace the evolution of research on *Big Data* applications in detecting accounting fraud. Its purpose is not merely descriptive but forward-looking: to identify which methodologies and *Artificial Intelligence* models are most effective, and to highlight emerging trends and knowledge gaps that can guide future inquiry.

The integration of *Big Data* and *Forensic Accounting* symbolizes a broader shift—from retrospective analysis to predictive insight. This fusion is redefining how fraud is detected, investigated, and prevented. Recent bibliometric studies confirm the growing academic focus on this intersection, particularly in the development of algorithmic fraud detection models and the identification of tax evasion through digital analytics (Thakkar, Datta, Bhadra, Barot, Purohit & Dabhade, 2024).

A bibliometric approach is especially valuable in this field because it enables researchers to map how ideas evolve and interact. By quantifying scientific production and visualising collaboration networks, it provides a dynamic picture of how disciplines such as accounting, computer science, and law converge. *Science mapping* techniques further enrich this perspective by revealing how research clusters are conceptually connected and where innovation is emerging.

Ultimately, the advantages of bibliometric analysis go beyond metrics. They lie in its ability to tell the story of a field in transformation: one where *Big Data* and *Artificial Intelligence* are not just tools but catalysts of change in the quest for financial integrity (Passas, 2024). Understanding this landscape is essential to designing future research that is both technologically advanced and ethically grounded—a goal that sits at the very heart of accounting in the digital era.

2. Design and Methodology

A bibliometric study was conducted using quantitative indicators and network analysis to explore scientific production, collaboration, and impact in the field of Big Data and accounting fraud detection. This method is appropriate for mapping emerging disciplines, as it enables visualisation of conceptual and social structures (Milán-García, Uribe-Toril, Ruiz-Real & de-Pablo-Valenciano, 2019). Bibliometric analysis has been used in similar studies within accounting and data science to identify knowledge evolution and research gaps (Rey-Martí, Ribeiro-Soriano & Palacios-Marqués, 2016; Mongeon & Paul-Hus, 2016).

However, to fully understand the potential and implications of the use of Big Data in this context, it is essential to examine and analyse the current state of research in this field. Figure 1 shows a schematic indication of the methodology applied.

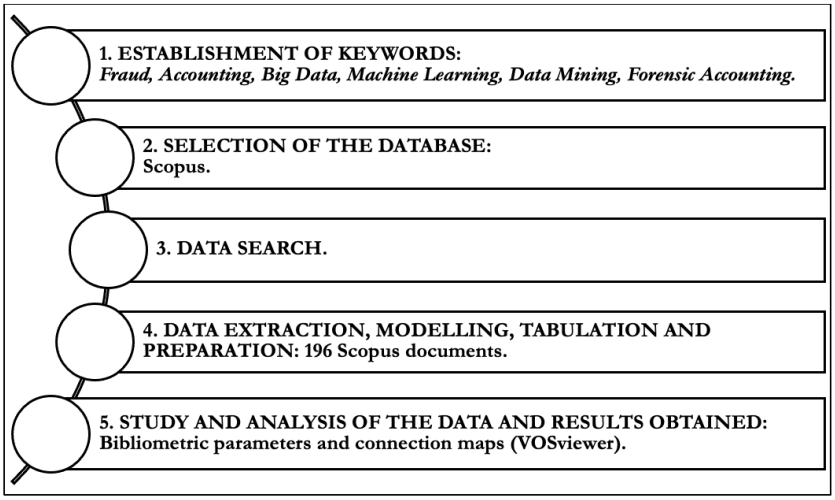


Figure 1. Methodological process of the bibliometric analysis (Milán-García et al., 2019)

2.1. Establishment of Keywords

The six keywords established for the study are Fraud, Accounting, Big Data, Machine Learning, Data Mining and Forensic Accounting. for this bibliometric analysis was a deliberate choice to build a well-rounded framework that captures the diverse and multidisciplinary nature of modern fraud detection in accounting systems.

Fraud has become a major concern in today's business world, with increasingly sophisticated schemes challenging traditional defences. At the heart of this fight is Accounting, which not only lays the groundwork for detecting fraud but also adapts to new technologies that both create risks and provide powerful tools to uncover suspicious activity.

Big Data plays a key role here, giving us the technological backbone to spot fraud by analysing enormous amounts of data for unusual patterns or signs of wrongdoing that would otherwise go unnoticed. Building on this, Machine Learning steps in with smart algorithms that learn from past cases, continuously refining their ability to catch fraud and outsmart schemes that change over time.

Then there's Data Mining, which turns raw financial data into valuable clues using techniques like clustering, classification, and association rules. But while technology is essential, it's the combination with human judgment and professional expertise that really makes the difference. Forensic Accounting brings it all together, blending cutting-edge tools with experienced eyes to not just spot fraud, but fully understand and investigate it.

The keywords "Fraud," "Accounting," and "Big Data" were chosen to define the study's conceptual scope, as they capture the core dimensions through which existing literature approaches technology-enabled fraud detection in accounting and forensic contexts.

Taken together, these keywords show that effective fraud detection today is an integrated effort. It relies on traditional accounting knowledge, advanced Data Analysis, and sophisticated technology, all working together. This selection of keywords ensures the study covers the full, interdisciplinary scope of modern fraud detection research, while staying grounded in practical solutions that help accounting professionals and organisations in the real world.

2.2. Selection of the Database

The selected database is Scopus, given its broad range of documents, including publications in all scientific disciplines and areas. It also enables users to examine data, such as citation analysis, by county, author, sources, and institutions (Rey-Martí et al., 2016).

Compared to other databases such as Web of Science (WoS), Scopus includes a higher number of journals in the fields of business, management, and information systems, which are central to the intersection of Big Data and accounting fraud detection (Rey-Martí et al., 2016). Several bibliometric studies have highlighted the greater representativeness of Scopus in emerging and multidisciplinary fields, making it particularly suitable for this research topic (Mongeon & Paul-Hus, 2016).

2.3. Data Search

The data collection process for this bibliometric study was carried out in May 2025 using the Scopus database. Specifically, the following search query was employed to retrieve relevant publications:

> TITLE-ABS-KEY (fraud AND accounting OR big AND data OR machine AND learning OR data AND mining OR forensic AND accounting).

This comprehensive search strategy aimed to encompass a broad range of studies at the intersection of Forensic Accounting, Big Data, Machine Learning, and Data Mining. Notably, during the data cleaning phase, no duplicated records were identified, nor was there a need to correct relevant errors in the dataset. Additionally, no weighting adjustments were made based on publication type (e.g., journal articles versus conference proceedings) and all authors were considered to maintain an unbiased representation of the scholarly landscape. Beyond this, descriptive data checks and initial data exploration confirmed the

consistency and reliability of the extracted dataset. These steps ensured the validity of subsequent bibliometric analyses and provided a fair foundation for mapping the evolution and impact of this interdisciplinary research area.

2.4. Data Extraction, Modelling, Tabulation, and Preparation

The data extracted were analysed and processed to obtain bibliometric parameters and indicators with respect to performance or productivity, as well as the impact of citations of the scientific output produced by the authors or researchers, countries, journals, and institutions. All the Scopus database records were exported in .csv format, so that they could be analysed statistically by Python and VOSviewer.

2.5. Study and Analysis of the Data and Results Obtained

Based on the 196 documents obtained in the search performed in the Scopus database and the complete extract contained in the Scopus .csv file, a statistical study of the main numerical variables was carried out with the Python programme (Bruce, Bruce & Gedeck, 2022).

For the numerical variables obtained, the describe ( ) function from the Python **pandas** library was used to return statistical information on the data gathered from Scopus. This method provides key insights, including sample count, mean, standard deviation, minimum, maximum, median, and the 25% and 75% percentiles, allowing for a comprehensive overview of the dataset's distribution and variability.

To create network and connection maps, VOSviewer software was used. This is a programme that supports bibliometric studies, making it possible to visualise structures and networks, in which labels and circles indicate frequency in the terms that correlate most strongly: the higher the frequency, the larger the circle or label (Arruda, Silva, Lessa, Proença & Bartholo, 2022). VOSviewer also enables users to change the colour of the groups or clusters, making it possible to analyse their constituent elements clearly, and establishes the distance, which is an indication of co-citation, co-authorship, and co-occurrence, as explained below:

- Co-citation is a research methodology in Library and Information Science (Rong, 2024). Its purpose is to identify the information flows or groups of authors or researchers that are established in the literature and to show the relationships between the nodes, which can represent the various actors involved (authors, journals, publications, etc.). It is therefore very different from simple citation counting, because a citation count is based on citations received from a third document.
- Co-authorship networks are used to analyse how the various actors are related to other research fields. This helps to detect groups of influential authors or researchers, hidden communities of authors, and collaborations, in general, between countries, institutions, etc. (Yan & Zhiping, 2023).
- Co-occurrence of words is understood to mean the appearance of two terms together in a given text corpus. This technique is used to analyse the content of a text from the joint occurrence of pairs of items, represented by terms or words, that make it possible to identity relationships between concepts within a particular domain. The greater the frequency with which the words appear together, the more strongly they are conceptually linked (Barcroft & Sunderman, 2023).

3. Outputs and Results

The findings are structured into eight subsections that highlight both quantitative and qualitative insights, starting with an exploration of country-level metrics like productivity, collaboration patterns, citation impact, and co-authorship dynamics. The analysis then shifts to an author-level perspective, using these same metrics to reveal how individual researchers are shaping and contributing to the field.

To bring depth and clarity to the results, Python has been used for statistical analyses and explored the most highly cited documents to identify key influences in the field. Temporal trends are unpacked through a year-by-year analysis of document distribution, while different publication types are compared to see how research is shared across formats.

The investigation also extends to sources of publications, examining journals, conferences, and other venues to reveal their roles in advancing this research area. Finally, a word cloud visualisation captures the most common themes and concepts, offering an intuitive view of the field's core ideas and emerging topics.

Together, these insights provide a comprehensive understanding of the research ecosystem, covering everything from productivity patterns to the big-picture themes shaping this domain.

The results obtained and the main outputs for the data analysed are shown below.

3.1. By Country

3.1.1. Productivity

For the 196 documents obtained from consulting Scopus, the following conclusions were obtained.

Country	Number of documents	%
United States	36	18.37%
India	18	9.18%
Indonesia	18	9.18%
China	14	7.14%
Malaysia	12	6.12%
Jordan	11	5.61%
Saudi Arabia	11	5.61%
South Korea	11	5.61%
Nigeria	7	3.57%
South Africa	7	3.57%
Italy	5	2.55%
Bahrain	4	2.04%
Germany	4	2.04%
Others	38	19.41%
Total	196	100.00%

Table 1. Most productive countries (compiled by the author from the data obtained from Scopus)

The first 13 countries with the most publications represent 80.59% of the total; the publication of documents is therefore highly concentrated in certain countries, among which the United States in first place with a total of 36 publications and then India and Indonesia with a total of 18 each.

It highlights how countries like the United States (18.37%), China, India, and South Korea are leading the way in using Big Data to tackle fraud. These countries have tech know-how, big financial sectors, and strong research programmes to build effective fraud detection tools. Interestingly, emerging economies like Indonesia, Malaysia, and some Middle Eastern nations also stand out, showing how the need for smart fraud detection is spreading as financial systems go digital around the world. In short, these patterns reveal where the world's biggest risks of fraud are meeting the most advanced solutions.

3.1.2. Collaboration

Collaboration between countries has also been analysed with the aid of VOSviewer software, selecting the countries with the highest number of citations and analysing their bibliographical couplings. Two groups of clusters can be seen: Indonesia and the United States (in red) and the India and Indonesia (in blue). Countries in the same cluster collaborate in some way, but there is also collaboration with the countries in the other clusters.

Based on these network visualisations, it's clear that some countries stand out for their strong collaborations and wide-reaching academic ties. The United States leads the way, closely working with partners like South Korea, Colombia, Serbia, and Canada. Meanwhile, India and Indonesia are at the heart of another busy network of cooperation, linking with Pakistan, Jordan, Vietnam, and Ghana. European and Middle Eastern countries like the

Netherlands, Greece, Slovakia, and Qatar also bridge these broader networks, fostering cross-regional exchange. Finally, smaller clusters such as Bahrain and Ecuador represent more niche collaborations, still vital in their specialized fields.

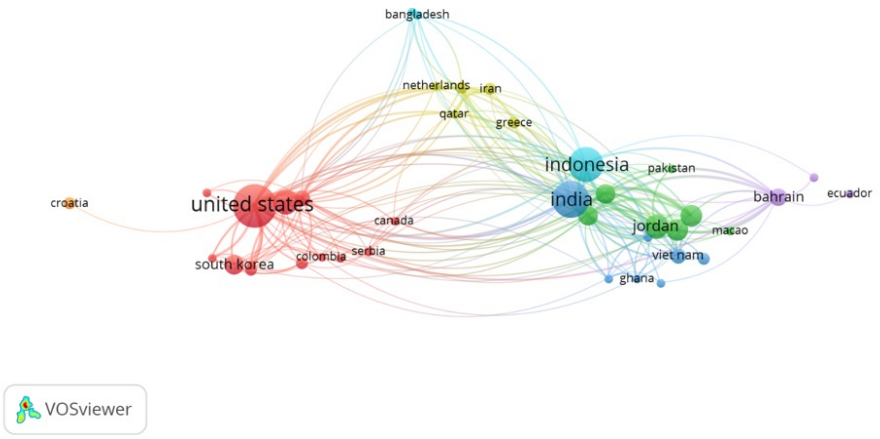


Figure 2. Visualisation map for bibliographical coupling by country (compiled by the author from VOSviewer)

3.1.3. Citations

The citations obtained from the Scopus search confirm, as indicated in Table 2, that the 7 countries with the most publications are also those that have the highest numbers of citations, accounting together for 68.77% of all citations.

The United States isn't just the country with the most documents in this study, it's also the most cited, with a share of citations even higher than its share of publications. India and China also stand out as countries with not only a high volume of research on financial fraud detection, but also as some of the most frequently cited after the United States.

Country	Number of citations	%
United States	1.001	38.98%
India	197	7.67%
China	128	4.98%
Qatar	128	4.98%
South Africa	118	4.60%
Malaysia	98	3.82%
Indonesia	96	3.74%
Others	802	31.23%
Total	2.568	100.00%

Table 2. Countries with the most citations (compiled by the author from the data obtained from Scopus)

3.1.4. Co-Authorship

Co-authorship of documents between countries was analysed, performing the search for all the countries that have published. Co-authorship was analysed by country, identifying 7 different clusters according to their proximity in co-authorship, as shown in Table 3.

Cluster 1 (9 items)	Cluster 3 (5 items)	Cluster 5 (1 item)
Egypt	Indonesia	Arab Open University
France	Malaysia	Cluster 6 (3 items)
Hungary	Oman	Nigeria
Iraq	Taiwan	South Africa
Jordan	Thailand	South Korea
Kuwait	**Cluster 4 (5 items)**	**Cluster 7 (3 items)**
Pakistan	Brazil	India
Palestine	California	Spain
Saudi Arabia	Hong Kong	United Arab Emirates
Cluster 2 (5 items)	Ireland	
Australia	United States	
China		
Iran		
Italy		
Switzerland		

Table 3. Clusters generated by the search for co-authorship by country (data obtained directly from consulting Scopus)

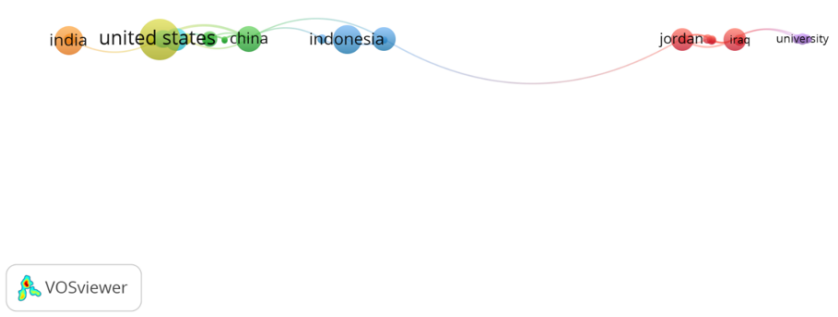


Figure 3. Visualisation map for co-authorship between countries (compiled by the author from VOSviewer)

Major hubs include countries like the United States, China, and Indonesia, while smaller clusters reflect regional partnerships or specialized institutions. Overall, it shows the global nature of academic collaboration today.

3.2. By Author

3.2.1. Productivity

Author productivity is indicated in Figure 4, where there are only five authors that have produced four publications conforming to the six selected keywords: Fraud, Accounting, Big Data, Machine Learning, Data Mining and Forensic Accounting. There is therefore no concentration for authors and their knowledge. There are also 7 studies for which the author's name is not available in Scopus.

3.2.2. Collaboration

Collaboration between countries has also been analysed with two maps for bibliographical coupling by authors, the first one with a maximum number of 4 authors per document and a minimum number of documents of one and, the second one with a maximum number of 1 author per document and a minimum number of one.

Despite the presence of some authors who publish individually, the analysis reveals that the most common and influential works are typically co-authored by small teams, often composed of three to four researchers. This pattern suggests a growing trend toward collaborative research, with the formation of stable and interconnected author networks.

Figure 4. Authors with the highest productivity (data obtained directly from consulting Scopus)

3.2.3. Citations

Using VOSviewer, a map has been created showing the most cited authors and their relationships.

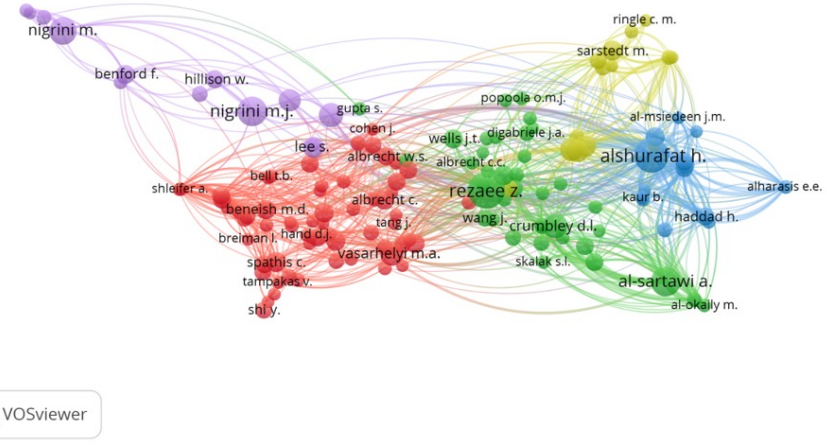


Figure 5. Visualisation map of the most cited authors (compiled by the author from VOSviewer)

The Table 5 shows the 18 authors with the most citations (taking co-authorship into account, since the documents are usually produced by more than one author), ordered from higher to lower numbers of citations for the documents they have published. The conclusion is that the authors with most citations are not those that have published the most documents, except for Lee, S. and Zerihun, M.F.; only two of the five authors who have published four documents are present in the table. Therefore, the authors with more documents obtained few citations.

3.2.4. Co-Authorship

A visualisation of co-authorship between authors, their interrelationship, and their link strength has been produced. The authors with the highest link strength for whom a visualisation has been produced in Figure 6 are not those with the highest number of citations except for Lee Sangjin. This can be confirmed by comparing Table 4 and Table 5.

Author	Citations	Total link strength
Rezaee, Z.	68	3.153
Alshurafat, H.	58	4.570
Al-sartawi, A.	43	3.136
Nigrini, M.J.	42	753
Nigrini, M.	37	504
Crumbley, D.L.	30	1.182
Durtschi, C.	29	830
Vasarhelyi, M.A.	28	1.992
Akinbowale, O.E.	26	1.427

Author	Citations	Total link strength
Zerihun, M.F.	25	1.395
Pacini, C.	25	696
Klingelhofer, H.E.	22	1.325
Lee, S.	22	219
Sarstedt, M.	21	1.405
Spathis, C.	20	2.491
Singleton, T.W.	20	780
Al shbail, M.O.	19	1.764
Zhang, J.	19	1.346

Table 4. Authors with the most citations for the documents they have published (data obtained directly from consulting Scopus)

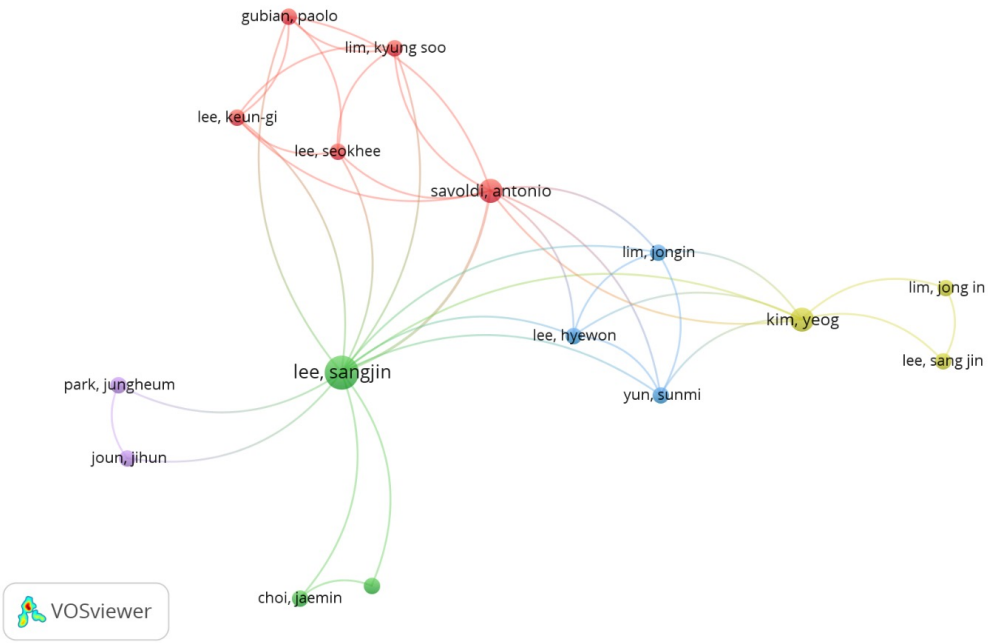


Figure 6. Visualisation map of co-authorship between authors (compiled by the author from VOSviewer)

Author	Docs.	Citations	Total link strength
Lee, Sangjin	4	27	14
Alrawashdeh, Badi	4	13	12
Savoldi, Antonio	2	9	10
Dabaghia, Mohammed Nadem	3	12	9
Darwazeh, Riyad Neman	3	12	9
Akinbowale, Oluwatoyin Esther	4	89	8
Brisimi, Theodora S.	1	1	8
Cucci, Fabrizio	1	1	8
Davis, John	1	1	8

Author	Docs.	Citations	Total link strength
Ferguson, Grace	1	1	8
Kristiansen, Morten	1	1	8
Lopez, Vanessa	1	1	8
Rho, Valentina	1	1	8
Scalvini, Jillian	1	1	8
Segrave-daly, John	1	1	8
Alharasis, Esraa Esam	2	10	7
Alkababji, Majdi Wael	1	0	7
Alzahrane, Mohammed	1	0	7

Table 5. Co-authorship between authors (data obtained directly from consulting Scopus)

Table 5 shows co-authorship data for 18 authors. Lee Sangjin is the most connected, while Oluwatoyin Esther Akinbowale has the most citations.

Three of the four authors who wrote the most cited document listed in Table 7 also appear in Table 5 with a link strength equal to or greater than 9. This suggests a strong correlation between academic influence (as measured by citations) and network centrality (as indicated by strength), reinforcing their key role in the field.

Many authors appear to have collaborated on the same paper, showing strong network links despite limited publications or citations.

3.3. Statistics Obtained from Python

A statistical summary of the numerical variables obtained from the Scopus.csv file of the Scopus database has been produced. The summary can be viewed in Table 6. Only the "Year" column and the "Cited By" column can be analysed. The other columns, in principle, are of no interest for this study.

The documents analysed were published between 2000 and 2025, with a median year of 2021, showing that most research in *Big Data* is recent and concentrated in the last decade. The standard deviation (5.35 years) confirms moderate dispersion, reflecting steady growth over time.

Citation impact is uneven: while the average or mean is 13 citations, the median is only 4, and a few papers reach over 200 citations, indicating that a small number of influential works drive much of the field's visibility. This pattern is typical of an emerging and fast-evolving academic area.

Statistical measure	Year	Cited by
count	196	196
mean	2.019	13
std	5.35	30
min	2.000	0
25%	2.016	1
50%	2.021	4
75%	2.023	10
max	2.025	247

Table 6. Summary of statistical data from the numerical columns of the Scopus.csv file obtained from the search (data obtained from the Python programme according to data extracted from Scopus)

3.4. Most Cited Documents

On the basis of the data obtained from Scopus, the 10 documents with the most citations have been placed in order, as shown in Table 7.

Position	Title	Year	No. of citations
1	Benford's law: Applications for forensic accounting, auditing, and fraud detection	2012	247
2	Leveraging Financial Social Media Data for Corporate Fraud Detection	2018	164
3	Data mining applications in accounting: A review of the literature and organizing framework	2017	128
4	Using machine learning to detect misstatements	2021	125
5	Accountants' perceptions regarding fraud detection and prevention methods	2006	122
6	FraudMiner: A novel credit card fraud detection model based on frequent itemset mining	2014	120
7	A comprehensive survey of data mining-based accounting-fraud detection research	2010	110
8	Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations	2012	78
9	Relevance of big data to forensic accounting practice and education	2019	75
10	Data mining journal entries for fraud detection: An exploratory study	2010	69

Table 7. Documents with the most citations (compiled by the author from the data obtained from Scopus)

As we can see in Table 7, the 10 most cited academic works in the field of *Forensic Accounting* and *fraud* detection from 2006 to 2021. These influential papers have significantly shaped the development of the discipline and offer insights into the main research trends over the past 15 years.

The studies mainly focus on three key areas: applying *Data Mining* and *Machine Learning* to detect *Fraud*, evolving Forensic *Accounting* techniques, and integrating technology into *accounting*.

Hot topics include Benford's Law, credit card *fraud*, and *Big Data* Analytics. The most cited paper, on using Benford's Law in audits and *fraud* detection, has 247 citations. Others explore social media analysis and Data Mining, showing a strong academic interest in tech-driven methods.

NIA-ES 240 "The auditor's responsibilities in an audit of financial statements with respect to *fraud*" (Instituto de Contabilidad y Auditoría de Cuentas–ICAC, 2024) emphasizes the auditor's responsibility to identify and assess risks of material misstatement due to *fraud*. Benford's Law can support this process by detecting anomalies in numerical data that may indicate manipulations or irregularities

Most high-impact papers were published between 2010 and 2018, but newer studies (2019-2021) also feature, reflecting ongoing innovation, especially in *Big Data* and Artificial Intelligence.

Overall, the field shows a clear shift from traditional approaches to more advanced, data-powered tools. With citation counts ranging from 69 to 247, these works have played a key role in shaping modern Forensic *Accounting*.

The authors who produced the documents are as follows:

Position	Authors
1	Al-Balqa' H.M.A.A.; Alrawashdeh B.; Dabaghia M.N.; Darwazeh R.N.
2	Hashemi S.K.; Mirtaheri S.L.; Greco S.
3	Faccia A.; Pandey V.; Banga C.
4	Shalhoob H.; Halawani B.; Alharbi M.; Babiker I.
5	Utama A.A.G.S.; Basuki B.
6	Ahmad V.; Goel R.; Arora M.; Venaik A.; Kumar R.
7	Gangwani M.
8	Archna R.; Bhagat N.
9	Kwak W.; Shi Y.
10	Haddad H.; Alharasis E.E.; Fraij J.; Al-Ramahi N.M.

Table 8. Names of the authors of the documents with the most citations
(compiled by the author from the data obtained from Scopus)

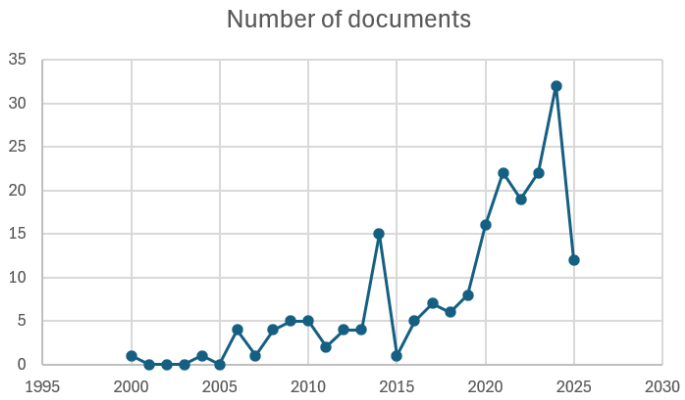
The authors of the most cited documents (except for the most cited; in the first position) are not those with the greatest importance, weight, and connections in Figure 5.

3.5. Documents by year

The year of publication of the documents obtained in the Scopus search was analysed.

If we draw up a table by periods of approximately five years, we can see that the documents are concentrated in the last five years, which account for 54.59% of all those published.

The data are therefore consistent with the novelty of the subject-matter addressed, not so much for the keywords *Accounting*, *Fraud and Forensic Accounting,* but for the increasingly intensive use of *Big Data, Machine Learning and Data Mining* given the computing capacity that has been attained in recent years and is expected to grow exponentially in the coming years.



Note: The data for 2025 corresponds to five months of the year.

Figure 7. Publications by year (data obtained directly from consulting Scopus)

Year	**Number of documents**
2021-2025(*)	107
2016-2020	42
2011-2015	26
2006-2010	19
2000-2005	2
Total	**196**

(*) The data for 2025 corresponds to the first five months of the year.

Table 9. Number of documents by period of years (compiled by the author from the data obtained from Scopus)

3.6. Types of documents

On the basis of the data exported from Scopus, the typology of the 196 documents obtained has been analysed.

Type	**Number**	**%**
Article	112	57.14%
Conference Paper	48	24.49%
Book Chapter	21	10.71%
Conference Review	7	3.57%
Book	5	2.55%
Review	2	1.02%
Letter	1	0.51%
Total	**196**	**100.00%**

Table 10. Types of documents (compiled by the author from the data obtained from Scopus).

Most of the documents are therefore articles and papers, within an academic field. The next section analyses the sources and the institutions that generated the documents.

3.7. Sources

A table has been compiled from the data obtained from Scopus to analyse the sources of the documents under consideration.

Most publications are clustered in sources related to Artificial Intelligence and computing, such as the ACM International Conference Proceeding Series or Studies in Systems Decision and Control.

This suggests a strong concentration of research around tech-driven approaches. In contrast, *accounting* and finance-related publications, like those in the Journal of Financial Crime or Issues in *Accounting* Education, are more scattered across various journals, reflecting a broader but less centralized academic interest. The high

number of documents classified as "Others" further highlights this dispersion in the financial and *accounting* fields.

If we analyse co-citation by different sources, we can obtain a visualisation of the connections, giving rise to various clusters.

Source	Number of documents
Journal of Financial Crime	9
ACM International Conference Proceeding Series	7
Issues in Accounting Education	5
Studies in Systems Decision and Control	5
Advances in Intelligent Systems and Computing	4
Cogent Business and Management	4
International Journal of Accounting Information Systems	3
Journal of Financial Reporting and Accounting	3
Lecture Notes in Computer Science Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics	3
Revista De Gestao Social E Ambiental	3
Wit Transactions on Information and Communication Technologies	3
Others	147
Total	**196**

Table 11. List of sources (compiled by the author from the data obtained from Scopus).

Source	Cluster	Citations	Total link strength
A guide to forensic accounting investigation	1	24	199
Accounting horizons	1	34	441
Accounting research journal	1	18	194
Critical perspectives on accounting	1	18	160
Issues in accounting education	1	54	305
J. bus. ethics	1	19	35
Journal of financial crime	1	80	679
Journal of financial reporting and accounting	1	22	210
Journal of forensic & investigative accounting	1	22	234
Journal of forensic and investigative accounting	1	28	413
Managerial auditing journal	1	85	1041
Research journal of finance and accounting	1	19	216
Sustainability	1	31	208
The cpa journal	1	23	262
Contemporary accounting research	2	37	924
Journal of accounting and economics	2	25	665
Journal of accounting research	2	42	882
Journal of financial economics	2	26	812
The accounting review	2	62	1306
The journal of finance	2	20	514
Auditing: a journal of practice & theory	3	31	560
Auditing: a journal of practice and theory (repeated)	3	27	297
Journal of accountancy	3	26	173
Journal of forensic accounting	3	44	347
Proceedings of the American philosophical society	3	19	149
Decision support systems	4	35	376
Expert systems with applications	4	30	177

Source	Cluster	Citations	Total link strength
International journal of accounting information systems	4	21	213
Journal of accounting and public policy	4	19	382
Journal of business ethics	4	26	455
Expert syst. appl.	5	41	367
Journal of emerging technologies in accounting	5	37	534

Table 12. Co-citation by different sources (data obtained directly from consulting Scopus)

Cluster 1 is shown in red, cluster 2 in green, cluster 3 in blue, cluster 4 in yellow and cluster 5 in purple.

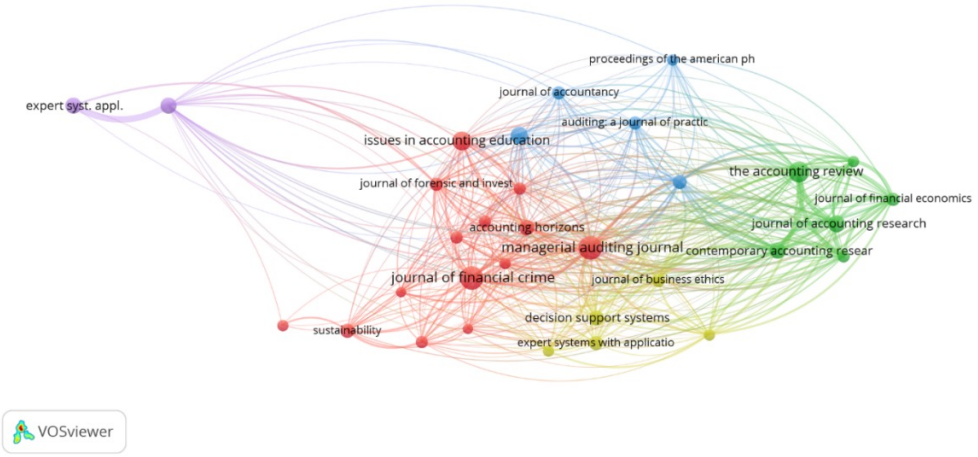


Figure 8. Visualisation map of co-citation between sources (compiled by the author from VOSviewer).

The co-citation map shows five distinct clusters of academic sources. Traditional and high-impact *accounting* journals form the core (e.g., The *accounting* review, Journal of *accounting* research), while others focus on *Forensic* A*ccounting*, education, ethics, and emerging technologies. This highlights the field's diversity, bridging classic research with tech and ethical concerns.

The sources with the strongest presence in terms of publication (like Journal of financial crime, Issues in *accounting* education, and Managerial auditing journal) stand out for their consistent focus on practical and emerging issues in *accounting*. Their high citation strength suggests they're widely recognized and influential within the field, often serving as key references in both academic and applied research.

3.8. Word cloud

A word cloud was created with the links for the volume of co-occurrence of all the keywords, with a minimum of occurrences of 3 keywords.

The map paints a clear picture: *accounting* research is evolving quickly, blending traditional topics like *fraud* and financial statements with powerful new tools from Artificial Intelligence and Data Science.

The keyword map shows how research in *accounting* and *fraud* is increasingly shaped by technology and Data Analysis. Five main clusters emerge:

- The red cluster groups terms related to Artificial Intelligence, highlighting its growing role in automating and enhancing *fraud* detection.
- The green cluster centers on *Big Data*, emphasizing its relevance in handling large volumes of financial and forensic information.

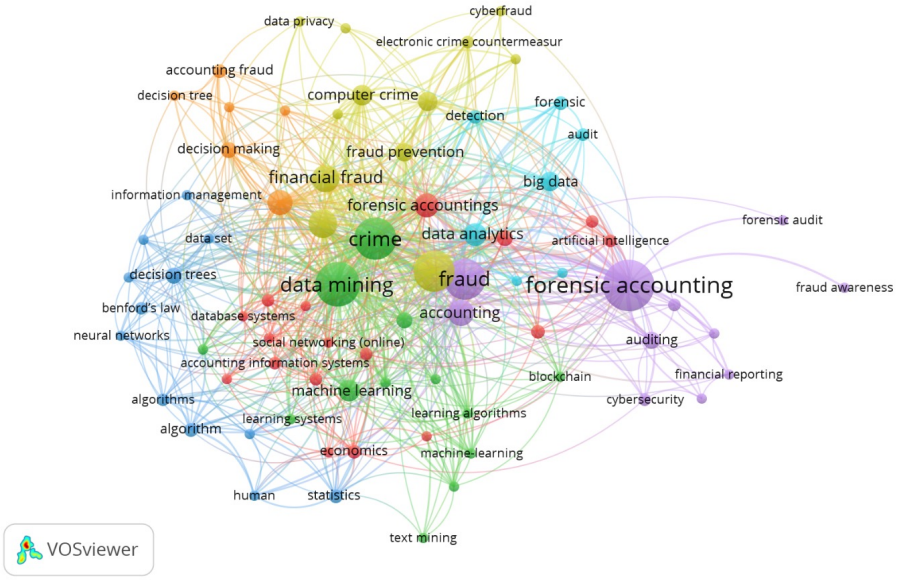


Figure 9. Visualisation map of all keywords (compiled by the author from VOSviewer).

- The purple cluster brings together Finance and Crime, pointing to the connection between financial systems and fraudulent activities.
- The blue cluster focuses on *Forensic Accounting* and *Fraud*, showing a strong interest in investigative and detection practices.
- The orange-yellow cluster includes Data Analytics, Financial Statements, and *Fraud* detection, reflecting practical tools used to ensure financial accuracy and integrity.

Figure 10 shows a chart that highlights key trends in *Forensic Accounting* research. Terms like *Forensic Accounting*, Crime and *Fraud* are most frequent among others, while tech-related keywords like *Machine Learning* show growing relevance.

As expected, the most frequent keywords are *Big Data* (9), Crime (42), *Fraud* (40), Finance (19), *Fraud* detection (41), *Forensic Accounting* (64), *Forensic Accountings* (14), *Accounting* (17), Financial statements (15), Financial Fraud (18), *Machine Learning* (11) and *Data Mining* (48). However, other frequent terms like Data Analytics, Computer crime and Digital forensics reinforce the trend toward digital and tech-driven approaches.

To complete this section, it offers a comprehensive bibliometric analysis of research at the intersection of *Big Data* and *accounting fraud* detection. The findings reveal an increasingly dynamic and interdisciplinary field that is gaining substantial momentum, particularly in the last five years. Over 54% of the analysed publications were produced between 2021 and 2025, signalling an accelerated interest in the application of data-driven technologies to combat financial misconduct.

A central insight from the study is the progressive shift from traditional audit and *Forensic Accounting* practices to more advanced technological approaches, particularly those leveraging *Machine Learning*, D*ata Mining*, and artificial intelligence. These tools are being applied not only to identify anomalies in financial datasets but also to develop real-time *fraud* detection systems capable of predictive analytics. The keyword co-occurrence analysis confirms the emergence of this trend, showing a convergence of terms like "*Forensic Accounting*," "*Fraud* detection," "*Big Data*," and "*Machine Learning*."

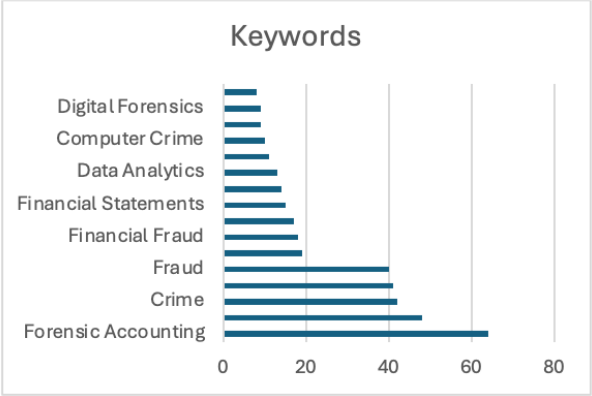


Figure 10. Keywords sorted by frequency (data obtained directly from consulting Scopus (Word cloud Figure 9))

From a geographical perspective, the production of scientific output is highly concentrated in technologically advanced economies such as the United States, China, and India. However, emerging economies like Indonesia and Malaysia are also increasingly contributing to this research area, reflecting a global acknowledgment of the importance of digital *fraud* detection as financial systems evolve. Collaborative networks across continents (especially between Asia, the Middle East, and North America) further reinforce the international scope of this field.

Despite this growth, the study also reveals notable discrepancies. For instance, there is little overlap between the most prolific and the most cited authors. This suggests that influence within the field is not merely a function of output volume but rather of the novelty and impact of individual contributions. Similarly, while certain journals such as Journal of Financial Crime and Managerial Auditing Journal serve as key publication venues, the dispersion of relevant research across numerous outlets, including those in computer science and information systems, reflects the inherently interdisciplinary nature of the domain.

A paradox also emerges while *Big Data* and Artificial Intelligence are heralded as the future of *fraud* detection, the very digitization of financial systems that enables these innovations also creates new vulnerabilities. In other words, the tools designed to detect *fraud* are evolving within the same technological ecosystem that facilitates increasingly sophisticated fraudulent schemes.

Looking ahead, the research points to several key trends. First, we can expect the continued integration of artificial intelligence in *accounting* and auditing, with a movement toward continuous, real-time monitoring of financial transactions. Second, future studies are likely to deepen the ethical and social dimensions of *fraud* detection, exploring how technology can not only identify wrongdoing but also foster a culture of transparency and accountability. Third, the field will likely fragment into specialized subdomains, focusing on sector-specific *fraud* risks (such as ESG reporting *fraud*, tax evasion, or cyber-enabled financial crime) each requiring tailored analytical methodologies.

Summarizing, the use of *Big Data* in *accounting fraud* detection is not merely a technological advancement but a strategic shift in how organisations and auditors approach financial integrity. As the field matures, the emphasis will likely move from exploratory to applied research, focusing on the implementation and regulation of these technologies in real-world contexts. A deeper understanding of interdisciplinary collaborations and methodological innovation will be essential for shaping the next decade of research in this critical area.

In short, as technology evolves and a deeper understanding of data is attained, it is reasonable to think that more intensive research will be conducted in the future. The use of *Big Data*, Data Analytics, and Artificial Intelligence will make it possible to improve the detection and prevention of *accounting fraud*, as well as the analysis of financial statements to obtain a more precise and accurate picture.

4. Discussion and Concluding Remarks

4.1. Originality and Main Findings

This study provides a bibliometric synthesis of research on Big Data applications in accounting fraud detection, positioned at the intersection of auditing, forensic accounting, and computational analytics. Its originality lies in combining multiple bibliometric lenses country productivity and citation impact, collaboration structures, author networks, publication formats, environments, and thematic clustering to explain not only how much the field has grown, but how it is structurally organising. In doing so, the paper offers an integrated interpretation of a domain that is increasingly shaped by data-driven approaches to fraud detection, while also revealing where conceptual integration with auditing frameworks and accountability requirements remains incomplete.

Bibliometric analyses suggest a research domain undergoing rapid consolidation, marked by growth that extends beyond volume to encompass structural development. Over half of the publications in the sample appear in the most recent window (2021-2025), consistent with the institutionalisation of data-driven fraud detection as a central agenda within accounting and auditing research. At the country level, research productivity and citation impact arises only partially. The United States, India, and China are prominent both in output and influence, yet the relationship between volume and impact is clearly imperfect. In particular, the United States accounts for a disproportionate share of citations relative to its publication count, a pattern consistent with cumulative-advantage dynamics in scholarly visibility, where hub status and network position may amplify recognition beyond what output alone would predict.

Collaboration structures further support this interpretation. International coupling and co-authorship networks suggest that central actors function as bridges across regions, facilitating knowledge diffusion and shaping agenda-setting. However, given the descriptive nature of bibliometric designs, these patterns should be interpreted as correlational regularities rather than causal mechanisms. Citation impact and collaboration intensity may reflect multiple confounders, including prestige hubs, language effects, disciplinary conventions, and preferential-attachment processes.

At the author level, the field exhibits a pronounced gap between productivity and intellectual influence. The limited overlap between the most prolific and the most cited authors indicates that impact is driven less by publication volume than by a smaller number of landmark contributions that define methodological toolkits and redirect research trajectories. Moreover, divergences suggest citation counts and network link strength suggests the coexistence of distinct influence profiles: "connectors" who occupy positions of relational centrality, and "high-impact specialists" whose work attracts outsized attention despite lower brokerage roles. Such duality is characteristic of emerging interdisciplinary domains in which methodological innovation can diffuse faster than conceptual consolidation.

The environment and keyword structures depict an inherently interdisciplinary ecosystem that remains only partially integrated. The substantial presence of conference papers alongside journal articles is consistent with fast iteration cycles in technology-oriented research streams, while co-citation patterns separate traditional accounting and auditing outlets from applied technology and decision-support platforms. In parallel, the keyword landscape foregrounds forensic and detection-oriented constructs (e.g., forensic accounting, data mining, fraud detection, digital forensics) more strongly than classical auditing concepts. This imbalance suggests that technological adoption is advancing more rapidly than its incorporation into auditing theory and professional judgement frameworks.

The bibliometric structure also highlights a set of tensions that deserve attention:

- The prominence of Machine Learning and analytics-oriented themes contrasts with the comparatively weak visibility of concepts associated with explainability and auditability, suggesting that technical performance has been prioritised more than interpretability.
- The strong representation of technology-oriented platforms and conference outputs, together with the relative absence of auditing-standard terminology in the thematic structure, indicates a gap between computational innovation and explicit engagement with auditing standards and evidence requirements.

- While Big Data and Artificial Intelligence are widely celebrated as enabling technologies, the keyword landscape provides limited evidence of sustained focus on bias, governance, and ethical safeguards, beyond higher-level reflections that tend to appear in concluding sections rather than in the main analytical core of studies.

Taking together, these tensions suggest that the field's methodological sophistication is advancing faster than its institutional and professional integration. Overall, the evidence supports a broader interpretation: the literature is shifting from rule-based auditing logics towards data-driven, probabilistic, and predictive epistemologies. The move from sample-based procedures to population level analytic enabled by computational method points towards technology mediated, potentially continuous assurance. Yet the bibliometric structures imply that this transformation is uneven: technological advances accelerate in applied streams, while integration with auditing theory, regulatory expectations, and accountability mechanisms lags. Big Data and Artificial Intelligence thus emerge not only as efficiency-enhancing tools, but as catalysts for structural change in how financial misconduct is conceptualised, detected, and potentially prevented.

Building on these results, the study makes three original, evidence-based contributions to the existing literature:

- It characterises a structural transition in fraud detection research. The combined evolution of publication trends, dispersion, and keyword clusters supports the interpretation of a shift towards technology-mediated assurance, where fraud risk assessment increasingly relies on population-level screening, anomaly detection, and predictive modelling rather than exclusively on periodic, sample-based procedures.
- It identifies an integration gap with direct relevance to audit practice. While the literature strongly emphasises Machine Learning and forensic analytics, the thematic structure suggests comparatively weaker attention to explainability, governance, and explicit alignment with auditing standards. From an auditor's perspective, this gap is material: models that are difficult to interpret, reproduce, or document may weaken audit trails, constrain defensibility, and reduce the usefulness of analytics as audit evidence. The study therefore reframes "advanced detection" as a dual requirement—analytical power and auditability.
- It offers a nuanced account of influence and diffusion in an interdisciplinary domain. By contrasting citation impact with network centrality, the paper distinguishes between "connectors" and "high-impact specialists", suggesting that the research agenda is shaped both by collaborative brokerage and by selective landmark contributions. This helps explain how methodological toolkits can diffuse rapidly even when conceptual consolidation remains incomplete.

4.2. Future Research Lines

The bibliometric patterns identified in this study suggest a research agenda that should move beyond technical performance towards stronger conceptual integration with auditing theory, evidence requirements, and accountability. Several lines of enquiry emerge directly from the structural features of the corpus of its thematic emphases and the uneven distribution of influence across countries and author networks.

Integrating analytics into auditing standards and evidence frameworks. The separation between technology-oriented outlets and traditional accounting and auditing platforms, together with the dominance of detection-focused terminology over classical audit constructs, indicates a persistent integration gap. Future research should examine how Machine Learning outputs can be translated into audit evidence in a manner consistent with audit documentation and evidential sufficiency. This includes clarifying how algorithmic signals inform risk assessment and materiality judgements, how professional scepticism is exercised when analytics mediate substantive testing, and how exception-handling protocols are designed to support defensible audit conclusions.

Explainable and auditable Artificial Intelligence for fraud detection. Although Machine Learning and advanced analytics occupy a central position in the thematic structure, concepts linked to interpretability, governance, and judgement appear comparatively marginal. Future studies should therefore prioritise the development and evaluation of explainable, auditable models that balance predictive accuracy with transparency. Key questions include the extent to which black box approaches can meet evidential transparency requirements, how model

documentation can be standardised for assurance contexts, and what forms of validation and reperformance are appropriate when detection relies on probabilistic outputs.

Bias, fairness, and model governance in high-stakes assurance settings. The limited visibility of ethics and governance-related themes suggests an underdeveloped agenda on bias and oversight. Research is needed to assess how algorithmic bias may affect false positives and false negatives across organisational, sectoral, or geographic contexts, and how governance mechanisms should mitigate these risks. This line of work should also address privacy considerations, accountability allocation, and escalation rules particularly where analytics influence investigative actions or control interventions.

Comparative and context sensitive research across regulatory and institutional environments. Country level concentration and network centrality patterns indicate that knowledge production is shaped by specific institutional settings. Comparative studies could examine how regulatory maturity, enforcement regimes, legal traditions, and professional norms condition the adoption, legitimacy, and performance of data-driven fraud detection. This would also help clarify transferability: whether models and detection approaches developed in highly digitalised economies generalise to contexts with different data infrastructures or fraud typologies.

Continuous assurance architectures and the human machine division of labour. The field's trajectory towards real-time or near-real-time monitoring raises foundational questions that remain insufficiently resolved. Future research should investigate organisational and technological requirements for continuous assurance systems, the appropriate division of labour between automated detection and human judgement, and the implications for auditor responsibility and liability when detection operates at scale and speed. Studies that link system design to governance and accountability arrangements would be especially valuable.

ESG-related misconduct and hybrid reporting assurance. Despite increasing regulatory and market emphasis on sustainability reporting, ESG-related manipulation is weakly represented in the thematic structure. This opens a timely frontier for applying Big Data and Artificial Intelligence to detect greenwashing, misreporting of impact metrics, and manipulation in sustainability disclosures. Research should also examine assurance implications in hybrid financial-non-financial reporting environments, including how evidence is obtained, validated, and documented for non-financial metrics.

4.3. Practical and Social Implications

The findings suggest implications that are directly relevant to auditing practice and to the public interest role of fraud detection.

From a professional auditing standpoint, the literature's emphasis on data mining, forensic analytics, and technology mediated detection supports the adoption of risk-based, analytics-enabled audit planning, where population level screening is used to prioritise journal entry testing and other high-risk procedures. However, the same evidence also implies that implementation must be accompanied by robust auditability safeguards: transparent documentation of data sources, modelling assumptions, thresholds, validation procedures, and exception-handling logic; clear protocols for re-performance; and explicit governance for how professional judgement corroborates or challenges algorithmic outputs. Without such safeguards, analytics may generate operationally compelling alerts but provide comparatively weak audit evidence.

The identified tensions further indicate a training and organisational imperative. Audit and control functions should invest in multidisciplinary capabilities, linking auditors, forensic specialists, and data scientists, so that analytics are embedded within the audit workflow (supporting risk assessment and the formation of audit conclusions) rather than treated as stand-alone technical exercises. In parallel, practitioners and standard-setters should prioritise guidance on explainable, governable, and ethically responsible Artificial Intelligence, addressing bias, privacy, accountability, and the defensibility of model-informed decisions.

Socially, the results reinforce that fraud detection is not only a technical optimisation problem but also a matter of trust and institutional integrity. As financial systems digitalise, the same infrastructures that enable more powerful detection can expand the opportunity set for sophisticated fraud, making governance and ethical oversight essential. Responsible deployment of Big Data and Artificial Intelligence can strengthen transparency,

reduce financial harm, and protect stakeholders provided it is aligned with fairness, privacy safeguards, and auditable accountability mechanisms.

4.4. Research Limitations

The selection of documents was incomplete, due to the fact that only the Scopus database was used. The incomplete selection is based exclusively on the use of this source of information and did not take account of any other factor or criterion. It is important to point out that this exclusion does not imply the consideration of any other resource or source of information in conducting the research study in question. This approach was adopted with the aim of maintaining the coherence and consistency of the methodology used.

The decision to use Scopus as the sole database for a bibliometric analysis is based on a combination of methodological reliability, comprehensive coverage, and consistency with previous scholarly practices. Scopus is recognized for its broad interdisciplinary coverage, indexing several peer-reviewed journals across all fields, including A*ccounting*, Finance, Computer Science, and Data Analytics; the core domains of this study.

The sample of six keywords inevitably conditions the scope of the results; alternative or expanded keyword sets could yield different thematic structures.

Finally, the predominance of English language publications may underrepresent contributions from non-Anglophone research communities.

These limitations should be considered when interpreting the findings and provide clear directions for future research, particularly through multi-database and multilingual approaches.

Declaration of Conflicting Interests

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author received no financial support for the research, authorship, and/or publication of this article.

References

Amat-Salas, O. (2017). *Empresas que mienten: Cómo maquillan las cuentas y cómo detectarlo a tiempo.* Barcelona: Editorial Profit.

Amat-Salas, O., & Lloret-Gual, P. (2020). Escándalos contables y crisis económicas. *Técnica Contable y Financiera,* 36, 18-24.

Arruda, H., Silva, E.R., Lessa, M., Proença, D. Jr, & Bartholo, R. (2022). VOSviewer and Bibliometrix. *Journal of the Medical Library Association,* 110(3), 392-395. https://doi.org/10.5195/jmla.2022.1434 (Accessed: October 2024).

Bao, Y., Hilary, G., & Ke, B. (2022). Artificial intelligence and fraud detection. In Babich, V., Birge, J.R., & Hilary, G. (Eds.), *Innovative technology at the interface of finance and operations* (11, 223-247). Springer Series in Supply Chain Management. Springer. https://doi.org/10.1007/978-3-030-75729-8_8

Barcroft, J., & Sunderman, G. (2023). *Lexis. In The Routledge handbook of applied linguistics* (218-232). Routledge. https://doi.org/10.4324/9781003082644-19

Bruce, P., Bruce, A., & Gedeck, P. (2022). *Estadística práctica para ciencia de datos con R y Python.* Madrid: Marcombo.

Chen, Y., Wu, Z., & Yan, H. (2022). A full population auditing method based on machine learning. *Sustainability,* 14, 17008. https://doi.org/10.3390/su142417008 (Accessed: October 2024).

Gay-de-Liébana-Saludas, J.M. (2013). Escándalos contables y financieros: de Banesto a Bankia. *Revista de Contabilidad y Dirección,* 16, 63-108. Available at: https://accid.org/es/revista-de-contabilidad-y-direccion-num-16/ (Accessed: October 2024).

Kayed, D., & Al-Sartawi, A. (2024). Forensic Accounting and Big Data a Literature Paper. In Al-Sartawi, A.M.A.M., Al-Qudah, A.A., & Shihadeh, F. (Eds.). *Artificial Intelligence-Augmented Digital Twins: Transforming Industrial Operations for Innovation and Sustainability, Studies in Systems, Decision and Control* (503, 303-312). Springer. https://doi.org/10.1007/978-3-031-43490-7_23

Instituto de Contabilidad y Auditoría de Cuentas (ICAC) (2024). Resolution of April 11, 2024, of the Institute of Accounting and Auditing, publishing the update of the Technical Auditing Standards (resulting from the adaptation of the International Standards on Auditing for application in Spain [NIA-ES] and other audit legislations). *Boletín Oficial del Estado,* 101.

Meiryani, M., Andini, V., Fahlevi, M., Yadiati, W., Purnomo, A., & Prajena, G. (2022). Analysis of accounting information systems based on artificial intelligence on fraudulent financial reporting trends in Indonesia. In *EBEE '22: Proceedings of the 2022 4th International Conference on E-Business and E-Commerce Engineering* (83-93). Association for Computing Machinery. https://doi.org/10.1145/3589860.3589871 (Accessed: October 2024).

Milán-García, J., Uribe-Toril, J., Ruiz-Real, J.L., & de-Pablo-Valenciano, J. (2019). Sustainable local development: An overview of the state of knowledge. *Resources,* 8(31). https://doi.org/10.3390/resources8010031 (Accessed: October 2024).

Mongeon, P., & Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: A comparative analysis. *Scientometrics,* 106, 213-228. https://doi.org/10.1007/s11192-015-1765-5

Passas, I. (2024). Bibliometric Analysis: The Main Steps. *Encyclopedia,* 4(2), 1014-1025. https://doi.org/10.3390/encyclopedia4020065 (Accessed: October 2024).

Rey-Martí, A., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2016). A bibliometric analysis of social entrepreneurship. *Journal of Business Research,* 69(5), 1651-1655. https://doi.org/10.1016/j.jbusres.2015.10.033 (Accessed: October 2024).

Rong, J. (2024). Co-citation and Visualized Analysis of Literatures in Library and Information Science Based on Web of Science and VOSviewer. *Academic Journal of Humanities & Social Sciences,* 7(8), 62-73. https://doi.org/10.25236/AJHSS.2024.070811

Saleh, I., Marei, Y., Ayoush, M., & Abu-Afifa, M.M. (2023). Big Data analytics and financial reporting quality: qualitative evidence from Canada. *Journal of Financial Reporting and Accounting,* 21(1), 83-104. https://doi.org/10.1108/JFRA-12-2021-0489 (Accessed: October 2024).

Sandhu, A.K. (2022). Big data with cloud computing: Discussions and challenges. *Big Data Mining and Analytics,* 5(1), 32-40. https://doi.org/10.26599/BDMA.2021.9020016

Selvarajan, G.P. (2021). Harnessing AI-Driven Data Mining for Predictive Insights: A Framework for Enhancing Decision-Making in Dynamic Data Environments. *International Journal of Creative Research Thoughts,* 9(2), 5476-5486.

Thakkar, H., Datta, S., Bhadra, P., Barot, H., Purohit, M., & Dabhade, S. (2024). A bibliometric analysis of forensic accounting research: Unveiling its impact on tax fraud detection in SAARC countries. *Journal of Informatics Education and Research,* 4(2), 21-22.

Yan, L., & Zhiping, W. (2023). Mapping the literature on academic publishing: A bibliometric analysis on WOS. *Sage Open,* 13(1). https://doi.org/10.1177/21582440231158562

Yang, M., & Zhang, J. (2023). Data anomaly detection in the internet of things: A review of current trends and research challenges. *International Journal of Advanced Computer Science and Applications,* 14(9). https://doi.org/10.14569/IJACSA.2023.0140901