How big data analytics is used in forensic accounting and auditing

Siriyama Kanthi Herath
Department of Accounting
School of Business Administration, Clark Atlanta University, USA

Angela Hamm
Senior Legal Assistant to the General Counsel
Clark Atlanta University, USA

Keywords
Big Data; Data Analytics; Forensic; Open Source, Accounting; Auditing

Abstract

Purpose

In this study, we investigate the application of big data analytics in forensic accounting and auditing. Big data technology mediates the connection between awareness and plans to use for fraud detection. We made an effort to assess how big data might be used in forensic accounting and auditing practices.

Design/methodology

This research is based on systematic literature review, and it provides a thorough understanding of the existing literature on big data research in forensic accounting and auditing. The research was gathered through peer-reviewed literature using current (post-2014) articles.

Results/findings

The model for corporate auditing has changed. Managers at the organization are proactively identifying and resolving internal abnormalities and discrepancies prior to an external audit using conventional audit techniques. Known white-collar crimes at several publicly traded corporations have led to a high demand for forensic accountants. Big data analytics is a key tool for spotting fraud patterns. Students should be encouraged to develop skills in data generation, sharing, analytics, ***mining, reporting, and storage because big data analytics is essential to business education.***

Practical implications and conclusions

The study's conclusions can be applied to the forensic accounting industry to assist in the development of Big Data fraud-fighting technology. Big data analytics and forensic accounting should be covered in accounting curricula. By offering different perspectives on big data and its practical implications in accounting practice, this research contributes to the body of published knowledge. As more companies use big data, the need to update the accounting curriculum is highlighted because success now requires a distinct set of skills. Future research should involve interviewing and surveying accountants, auditors, and other members of the financial services industry, which will be difficult but necessary.

1) The Introduction

In this study, we investigate the application of big data analytics in forensic accounting and auditing. Big data technology mediates the connection between awareness and plans to use for fraud detection. We made an effort to assess how big data might be used in forensic accounting and auditing practices. In order to foster understanding of modern big data techniques' potential applications, we first introduce them. It talks about the advantages and restrictions of using big data in forensic accounting and auditing activities.

Big data analytics is the act of dissecting huge amounts of data to find information that could aid businesses in making well-informed decisions about their operations, such as occult patterns, relationships, market dynamics, and customer preferences. Advanced analytics, which also includes elements like mathematical analyses, what-if analysis, and modeling techniques backed by analytics tools,

is a subset of big data analytics. The study's conclusions can be applied to the forensic accounting industry to assist in the development of Big Data fraud-fighting technology. We were compelled to learn more about the impacts of big data in order to advance our forensic accounting and auditing initiatives in the future.

This paper is structured as follows. The first portion defines the big data literature review, its frameworks, and its categorizations. The types of data analytics and the challenges associated with their use are explored in the second part. In the third section, the research methodology is outlined. In the fourth section, research findings are covered along with a detailed discussion of how Big Data Analytics is used in forensic accounting and auditing. In the fifth section, discussions and conclusions are offered. The final section includes limitations and suggestions for additional research.

2) Literature review

Both literally and symbolically, big data is "big." It is made up of bigger, more complicated data sets than can be handled by conventional processing tools. Big Data cannot be defined with absolute certainty. Big data simply refers to a sizeable amount of data that is difficult to process using conventional database and software techniques (Greer, 2013). Big data today includes both organized and unstructured data. Structured data is information that has already been arranged, is often numerical, and is easy to format, such as Excel spreadsheets or Access databases. Unstructured data lacks organization is in a free-form format, and is not arranged according to a predetermined model, like emails or films. Moreover, big data frameworks should enable both organized and unstructured data, distributed data programming, and high throughput (Singh & Reddy. 2014).

Integration, administration, and analysis are the three critical processes that make big data work. The ability to evaluate large data sets, such as terabytes and petabytes, is necessary for integration. Management includes selecting a data storage method, like the Cloud. Acting on the data found is the analysis. The following are some best practices for creating a solid Big Data foundation:

- Align with specific business goals.
- Assess skills early and often.
- Share knowledge across the enterprise; and
- Integrate unstructured with structured data (www.oracle.com).

2.1) Big Data Characteristics

Many characteristics or attributes that are listed by nV's characteristics are present in big data (Herath & Joshi, 2023, p.31). One of the first definitions was the 3Vs concept, which considered volume, velocities, and variety (Laney, 2012). This concept has evolved into the 4Vs with the addition of Veracity (Vossen, 2013). The concept has been further developed to include characteristics like volatility, viability, and variability. It is critical to remember that the most important V for the corporate environment is Value (Marchand-Maillet et al., 2014). For our study, we employ the 6Vs definition, which stands for Volume, Velocity, Variety, Value, Variability, Veracity, and Validity (Figure 1). Data can be characterized as "big" or "small" using the "6 V's":

1. Volume can be defined as the quantity of data or the size of a sizable collection.
2. Data processing and penetration speed is referred to as velocity. Enormous data is quickly gathered.
3. Data that can come from both internal and external sources might be referred to as variety. Data can be structured, semi-structured, or unstructured in terms of organization (unorganized).
4. Data's utility determines its value. Big data is valuable if it is useful.
5. Data variability refers to the dynamic nature of the data.
6. Veracity is the correctness and dependability of facts.

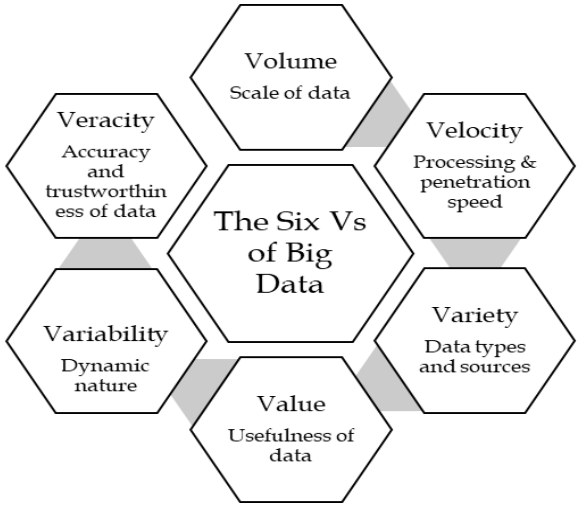

Figure 1: The Six Vs of Big Data

2.2) What is Big Data Analytics?

The phrases "big data" and "data analysis" are commonly used interchangeably. Big data analytics is the examination of vast amounts of saved data with the goal of identifying patterns of behavior. To identify a pattern, the data is gathered, saved, processed, and examined. As such big data analytics is the process of recognizing trends, trends, and correlations in immense amounts of unstructured data in order to assist data-driven decision-making. These methods employ well-known statistical analysis techniques, such clustering and regression, to larger datasets with the aid of more sophisticated technology.

2.2.1) Types of Analytics

Data analytics can be divided into four categories:

- Descriptive: Descriptive analytics use graphs or reports to describe historical trends. By using historical data to identify trends and patterns, knowledge can be achieved.
- Diagnostic - ascertain the cause of the incident. Diagnostic analytics assist businesses in identifying the root cause of an incident. Analysis of data help identify a problem's underlying causes.
- Predictive - foresee potential outcomes (most useful for companies). Predictive analytics uses statistical models to forecast and make predictions.
- Prescriptive - provide explicit recommendations on the best course of action. Prescriptive analytics makes explicit recommendations and uses models to describe the optimal course of action.

Big data analytics tools and technologies can help businesses make data-driven decisions that will enhance the results of their business operations. Improvements in operational performance, marketing effectiveness, and consumer personalization are possible benefits.

3) Research methodology

This report provides a review of the literature on the application of big data techniques in forensic accounting and auditing research and practice. To determine the state of the field, we evaluated the prior research on big data in forensic accounting and auditing. The best strategy for this research area is a literature review because a company's website frequently lacks the necessary data. Printed documents for corporations don't frequently provide in-depth assessments. We used the most readily available information found in peer-reviewed journals to conduct our study. Finding out how organizations use big data in routine operations involved reviewing recent (post-2014) papers.

4) Findings/results

4.1) Big Data Analytics and Decision Making

Both large and small firms can benefit from the techniques used in big data analytics. Data visualization employs infographics and statistics to convey ideas. Patterns can guide decisions with the least amount of human input thanks to machine learning techniques like the neural network. Other programs, like Hadoop, enable big data analytics by generating queries. The Securities and Exchange Commission (SEC) employs RobotCop as a tool to identify financial statement problems and regulatory violations. Big Data is useful for tax data analysis as well. The New York Department of Revenue and Finance uses a fraud detection system to find income tax fraud (Dewu and Barghathi, 2019). Some nations have passed anti-money laundering laws requiring meticulous forensic accounting. Big data analytics makes it simpler to identify trends in fraud. Financial regulatory and compliance analytics help financial service businesses with standards for risk, behavior, and transparency in the auditing space ([1]https: www.oracle.com/big-data).

Testing of journal vouchers is fully automated. JVcATS enables for the input of trial balance data and performs journal entry testing automatically. Other sorts of software utilized include DacEasy and Aura (fraud detection). (Dewu and Barghathi, 2019). Data mining, semantic analytics, data visualization, geographic analysis, and data discovery are some methods used to extract knowledge from extracted data. To examine data in text, such as emails and tweets, techniques like text mining and semantics are helpful. Algorithms, natural language processing, and pattern recognition are more sophisticated techniques. Techniques such as artificial intelligence use decision trees and machine learning. Without much or any human input, machine learning develops patterns to help make decisions (Dewu and Barghathi, 2019). The Association of Chartered Certified Accountants (ACCA), a global organization, predicts that the impact of big data on standards will stop standards from becoming outdated (Dewu and Barghathi, 2019).

Data governance controls who have access to what kinds of data as well as which data types are governed. Data governance is the administration of data assets through control, authority, and group decision-making (Lin, 2014). When there are chances to increase efficiency, key performance indicators are defined and monitored, driver-based forecasting is employed, and external threats are monitored, better data quality is most beneficial (Lin, 2014). Leading data governance techniques include:

- A powerful chief officer position to ensure that data is correctly maintained and organized.
- Develop a plan that includes all relevant parties. This may be facilitated by clear roles and responsibilities, education, and ongoing review.
- An inventory and catalog of current and prospective data assets.
- Maintain control over data throughout its life. Establish stage management policies and procedures.

Below are a few methods for improving data quality:

- Examine the data's source. Identify the source and the stakeholder before interacting. Recognize data flows.
- Stress the value of excellence. Staff should be encouraged to identify the data that directly affect analysis.
- *Use technology to promote quality. Use an organized list of alternatives and built-in edit checks to locate and fix problems.*
- Examine the data entry procedures. Compare system data to source records to test the system on a regular basis.
- Connect data to other sets of data. In comparison with a related data set.
- Work with a third party. If corporate experience is lacking, seek advice from a respected outside public accounting or management firm (Lewis, Steinhoff and Chawda, 2017).

4.2) How Big Data Analytics Is Used in Forensic Accounting and Auditing

The Latin word forum is the source of the word "forensic." A forum in the context of business is a group of people who decide utilizing information presented by an accountant (s). Data mining and visualization tools are used by forensic accountants to close the gap between what is believed to be true

and what the evidence demonstrates to be reality. There are established laws regarding fraud. For publicly traded firms, the Sarbanes-Oxley Act (SOX) imposed financial and auditing rules. By enforcing these standards, the public and investors are shielded from fraud and accounting mistakes. The security and privacy of patient information are safeguarded by the Health Insurance Portability and Accountability Act (HIPAA). The GrammLeach-Bliley Act (GLB) safeguards the financial privacy of customers. The Act governs the gathering and disclosure of information, offers protections, and forbids the gathering of information using pretext (Rechtman, 2020).

Big data is crucial since fraud costs businesses 5% of their yearly revenue. 79% of firms use more than 10 million records, and they seek tailored solutions, according to a 2018 Ernst & Young survey. According to comparative surveys done in 2018, data analytics are being used more frequently in forensic/investigative services. Also revealed to be the fastest-growing fraud risks are insider threats and cyber breaches. Data visualization, predictive analytics, behavior analytics, content analytics, social network analysis, and geospatial analytics are all said to be used in day-to-day operations by respondents (Rezaee and Wang, 2019).

What can be done with big data? It can spot fraud in financial transactions, such as when a client deposits a lot of money and then withdraws it the following business day. Businesses can use big data to examine enormous volumes of data to identify patterns, stop disease, and fight crime (Debenham, 2016). When handling massive amounts of data, spreadsheet software performs worse than big data software tools. A few examples of analytical software are Tableau, MS Power BI, and Qlik. The software is easy to use, uses a variety of data, can integrate data from different sources, and can filter and clean data (Pan and Blankley, 2018).

The use of big data is prevalent in Deloitte's operations. It is used to increase efficiency and improve risk management. Data modeling has replaced repetitive work. Big data has made it possible for the company to offer more specialized counsel and financial planning to its clients (Chien, 2020). Data models are replacing monotonous activities at accounting firms. One benefit that endures for a long time is profitability.

4.3) Challenges with big data analytics

Despite the extensive benefits of embracing big data analytics, there are challenges as well: Big Data analytics has a number of challenges:

- Accessibility - As data volumes increase, processing and storage become increasingly difficult.
- Upkeep of data quality - Data come in a range of sources and forms. It takes a lot of time, energy, and resources to maintain effectively.
- Data security faces special security concerns in complex systems. For instance, many universities utilize the Banner system to handle transaction processing and manage student records. To avoid intrusions and the misuse of information, users must utilize two-factor authentication.
- Selecting the Correct Tools - Businesses must select the most appropriate tool to match user requirements and infrastructure (http: //www.searchbusinessanalytics.echtarget.com)

Traditional data processing methods cannot capture or interpret massive data due to its diversity and speed. The great thing about big data is that we can use the plethora of free open-source big data platforms to take advantage of it. These platforms are the combination of hardware infrastructures and software tools used to acquire, store and analyze data (Almeida & Bernardino, 2015, p.3). Open-Source platforms provide free solutions that are also simpler to customize to the unique needs of each enterprise. Cassandra, Hadoop, and Spark are the three main open-source frameworks used to store and handle big data. A distributed file system is used by Hadoop (DFS). Data is broken up into smaller pieces and kept in several machines. Hence, even if data is lost on one system, it will still be secure on another.

5) Discussions and conclusions

Big data helps accountants in intricate areas like risk management and fraud detection. By employing big data and their knowledge of business processes, CPAs may transform insight into competitive advantages. With big data, accountants assist organizations in concentrating on measurable outcomes like return on investment (ROI) (Lin, 2014). Deloitte has developed business analytics methods for its clients that increase efficiency and improve risk management.

The Sarbanes-Oxley Act (SOX) and the Health Insurance Portability and Accessibility Act are two examples of laws that have been built around the fundamental notion of fraud (HIPAA). The collapse of publicly traded businesses like Enron and WorldCom led to more stringent auditing standards and a growth in the number of CPAs who specialize in detecting fraud. Forensic accountants are experts at gathering, examining, and summarizing financial information for a decision-making body or individual. Business valuation, electronic discovery, due diligence, and securities and tax fraud are a few forensic accounting subfields. Big data is utilized in presentations that include simulations and case studies. Forensic accountants must have a realistic understanding of data set management and visualization methods (Rechtman, 2020).

The usage of forensic data analytics in forensic and investigative services has increased, according to results of surveys conducted by Ernst & Young in 2018. Instead of using conventional database approaches like query design, forensic accountants use geospatial analytics, predictive analytics, and data visualization. The Securities and Exchange Commission (SEC) uses "RobotCop"-designed big data analytics to find legal infractions and abnormalities in financial statements. (Rezaee and Wang, 2019). Leading businesses have security measures in place and manage their data assets, including the number, quality, and usage techniques. A leading practice is aware that data is gathered, arranged, used, managed, and destroyed throughout its lifecycle. The auditing paradigm has changed. Prior to external audits, firm managers utilize internal standard audit methods like data mining to find and fix irregularities (Lewis, Steinhoff and Chawda, 2017).

Big data is used in the financial services sector to spot fraud-detection patterns and streamline regulatory reporting. Big data helps businesses with compliance, risk, behavior, and transparency (www.oracle.com/big-data, 2020), which is another advantage. Deloitte employs software to carry out repetitive and backward-looking operations. Instead of analyzing a sample, entire sets of translation are examined. The big data anti-fraud model from Deloitte finds unusual occurrences like manipulation and money theft (Chien, 2020). Scatterplots and algorithms are used to find anomalies. Instead of using keyword searches to process eDocuments, big data is utilized. Big data analytics is used by anti-fraud specialists to monitor the online activity and behavior of fraudsters in order to anticipate the legitimacy of transactions in real time, such as credit card fraud. Real Time Big Data Architectures (RTBDA), which include four layers (data, analysis, integration, and decision), are used to accomplish this (Debenham, 2016).

Accountants and auditors are trailing behind in big data analytics. Interactive learning like Tableau, MS Power BI, and Qlik are frequently used to close the knowledge gap. The use of Tableau in teaching accounting analytics has been proposed. Accounting professionals can master the fundamental methods for analyzing big data by using Tableau (Pan and Blankley, 2018). Academic research is underdeveloped as a result of the big data industry's rapid growth. There is a large overlap between big data and audit evidence in terms of sources and forms. More resources and simultaneous, real-time auditing of the goods can uncover hazards for forensic accounting and fraud detection. Higher assurance and auditing quality levels are a big data advantage (Ibrahim, Elamer and Ezat, 2021).

For careers in the field, extensive big data analytical abilities are essential (Dewu and Barghathi, 2019). Business schools should incorporate forensic accounting and big data analytics into their curricula since fraud costs companies 5% of their annual revenue. Big data analytics, according to the Association for Advancement of Collegiate Schools of Business (AACSB), is essential to business education and promotes the growth of expertise in data generation, sharing, analytics, mining, reporting, and storage.

5.1) How Big Data Can Improve Education and Research

Big data is viewed as an asset on the financial sheet by 20% of large firms, according to the Association of Chartered Certified Accountants (ACCA). Big data will affect accounting standards by 2024. However, some accounting curricula don't teach analytical skills (Dewu and Barghathi, 2019).

Students should receive hands-on training in big data analytics as part of a distinct course that is incorporated in the curriculum. Human intuition and professional judgment are necessary skills. Core student competences include the use of technology and software for data extraction and analysis, the capacity for analytical thought and decision-making, and interpersonal abilities like problem-solving, organizing, and planning that lead to effective analysis. Rawls College of Business and Texas Tech University are two institutions that teach big data. The Greehey School of Business at St. Mary's University now offers a new curriculum called Accounting and Data Analytics (Dewu and Barghathi, 2019). Big data will aid forensic accounting students in developing their knowledge in the following areas:

- Perform data mining and modeling.
- Advanced analytical and data management skills.
- Extract, transform and leverage data.
- Clear database digital strategy (Rezaee and Wang, 2019).

6) Limitations and direction for future research

This paper aims to determine how big data analytics is used in forensic accounting and auditing. It discusses the benefit, and limitations of big data in the operations of forensic in this study, we explore how big data analytics can be used in forensic accounting and auditing. It discusses the benefits and limitations of utilizing big data for forensic accounting and auditing tasks. The goal of forensic accountants and auditors is to gather information. The types, definitions, and common uses of big data in the auditing and forensic accounting sectors were discussed. Businesses are using big data more and more. The study highlights how important it is to adapt the accounting curriculum to account for the new set of success-related skills.

The model for corporate auditing has changed. Managers at the organization are proactively identifying and resolving internal abnormalities and discrepancies prior to an external audit using conventional audit techniques. Known white-collar crimes at several publicly traded corporations have led to a high demand for forensic accountants. Big data analytics is a key tool for spotting fraud patterns.

Peer review was conducted on the literature that was used in the study. The study looked at recent (post-2014) publications to determine how companies use big data in everyday operations. This study is limited since data from firms must be acquired for a thorough empirical analysis. Because big data analytics is crucial to business education, students should be encouraged to enhance their skills in data collection, sharing, analytics, mining, reporting, and storage. Interviewing and surveying accountants, auditors, and other professionals in the financial services sector are important yet challenging tasks for future research. There should be more peer-reviewed studies and accounting journals examined for current understanding.

References

Almeida, Pedro, and Jorge Bernardino. "A Comprehensive Overview of Open-Source Big Data Platforms and Frameworks." Services Transactions on Big Data, vol. 2, no. 3, Oct. 2015, pp. 15–33.

Chien, H.-Y. (2020). Use of Business Analytics in Accounting Firms—Taking Deloitte as an Example. E3S Web of Conferences, 218, p.03004. doi:https://doi.org/10.1051/e3sconf/202021803004.

Debenham, David. (2016), "Big data analytics, big financial institutions, and big money fraud litigation" Banking and Finance Law Reviewer, pp.103, 143.

Dewu, K. and Barghathi, Y. (2019). The accounting curriculum and the emergence of Big Data. Journal of Accounting and Management Information Systems, 18(3), pp.417–442. doi:https://doi.org/10.24818/jamis.2019.03006.

Dynatrace. (n.d.). Business Analytics. [online] Available at:
https://www.dynatrace.com/monitoring/platform/business-analytics/?utm_source=google&utm_medium=cpc&utm_term=business%20analytics%20platform&utm_campaign=us-dem-

dem&utm_content=none&gclsrc=aw.ds&gclid=EAIaIQobChMIv9Ougdun_QIVghPUAR2mMAFvEAAYASAAEgKe2fD_BwE [Accessed 21 Feb. 2023].

Ernst & Young (2018) How can you disrupt risk in an era of digital transformation? Global Forensic Data Analytics Survey 2018 https://assets.ey.com/content/dam/ey-sites/ey-com/en_gl/topics/assurance/assurance-pdfs/ey-global-fda-survay.pdf

Greer Jr., M. B. (2013). 21st Century Leadership: Harnessing Innovation, Accelerating Business Success. iUniverse.

Ibrahim, A.E.A., Elamer, A.A. and Ezat, A.N. (2021) 'The convergence of big data and accounting: innovative research opportunities', Technological Forecasting and Social Change, 173, 121171, pp. 1-13. doi: 10.1016/j.techfore.2021.121171.

Herath, S. K., & Prem Lal Joshi (2023). Audit Data Analytics: A Game Changer for Audit Firms. International Journal of Auditing and Accounting Studies. Special Issue on Auditing in the Digital Age, 5(1), pp. 29-48. https://DOI: 10.47509/IJAAS.2023.v05i01.02

Laney D. (2012). 3D Data Management: Controlling Data Volume, Velocity and Variety. File 494.Gartner. http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Lewis, A., Steinhoff, J., and Chawda, V. (2017) "Overwhelmed by data? Using and protecting the right data", Journal of Government Financial Management, Fall 2017

Lin, P., (2014), "What CPAs need to know about big data", The CPA Journal, November 2014

Marchand-Maillet, S., Hofreiter, B. (2014) Big Data Management and Analysis for Business Informatics. Enterprise Modelling and Information Systems Architectures (EMISA), 9(1), pp.90-105.

Oracle (2021). What is big data? [online] www.oracle.com. Available at: https://www.oracle.com/big-data/what-is-big-data/.

Pan, K., and Blankley, A. (2018) "Using visualization software to compile and analyze data: A step by step guide for CPAs", The CPA Journal, June 2018. https://www.cpajournal.com/2018/06/27/using-visualization-software-to-compile-and-analyze-data/

Rechtman, Yigal M. (2020). "The past, present and future of forensic accounting", The CPA Journal, March 2020, pp. 10-11

Rezaee, Z. and Wang, J. (2019), "Relevance of big data to forensic accounting practice and education", Managerial Auditing Journal, Vol. 34 No. 3, pp. 268-288. https://doi.org/10.1108/MAJ-08-2017-1633

Singh, D., & Reddy, C. (2014). A survey on platforms for big data analytics. Journal of Big Data, 2 (8), pp.1-20.

Vossen, G. (2014), Big data as the new enabler in business and other intelligence. Vietnam Journal of Computer Science, 1(1), pp.3-14.