

Review Article

The Role of Big Data Technology in Revolutionizing the Insurance Industry in Kenya

Cyprian Omenge Nyambane* 

Department of Management Science, Kenyatta University, Nairobi, Kenya

Abstract

Recent years have witnessed the emergence of novel ideas and concepts of big data to face the remarkable rise of amounts of data in many business sectors. Meanwhile, the remarkable growth of internet use and social networks have even added not only huge amounts of data to different business sectors but have also added challenges to conventional data processing systems. To deal with the large amount of data, traditional processing techniques have proved to be inefficient and insufficient to provide accurate and meaningful information required for evidence-based decision making. The insurance industry has also heavily relied on processed data for accurate risk assessment, underwriting and pricing. However, the sector today is also dealing with huge amounts of accumulated data, both structured and unstructured, which has made traditional data processing techniques unable to handle. Big data Analytics is an aspect of innovation which has recently gained major attention from both academics and practitioners. Big data analytics is the process of examining big data to uncover information such as hidden patterns, correlations, market trends and customer preferences. These revelations can significantly impact an organization as it provides deeper insights into customer behaviour, operational efficiency, and market trends. This paper aims to assess the role of Big Data adoption in the insurance industry through literature review. The paper presents big data and insurance industry in Kenya, characteristics of big data, the technologies used in big data implementation. It also looks at the beneficial role of adopting big data technology in the insurance sector and concludes that big data analytics stands out as an enabler to insurance organisations in making data-based decisions and providing customized insurance products and services according to customer needs. Finally, the paper encourages future research to examine the levels of big data adoption and its effect in the insurance industry.

Keywords

Big Data, Cloud Computing, Artificial Intelligence, Internet of Things, Insurance Industry, Machine Learning Language

1. Introduction

In a data-fueled age, the insurance industry stands at a transformative crossroads. The insurance industry, for a long time, has been known for leveraging traditional business models. The industry continued its legacy business and products for quite some time. But with the intervention of modern-day technologies, the industry has witnessed expo-

nential growth of the use of technology like any other sector. Advanced technologies and digital platforms have allowed insurance companies to try new means of tracking, measuring, and controlling risk. Advanced technologies like artificial intelligence (AI), machine learning, and internet of things (IoT) have gained enormous popularity over the past few

*Corresponding author: cnyambane@gmail.com (Cyprian Omenge Nyambane)

Received: 30 April 2025; **Accepted:** 27 June 2025; **Published:** 8 September 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

years. Applications of AI such as chatbots, automation, machine learning, and predictive analysis have made it possible for computers or programs to comprehend and evaluate big data.

Competition, disruptive technologies, innovations, creativity and strategic marketing have all contributed to the rapid change in the insurance industry's current landscape. In order to sustain competitiveness in the current uncertain environment, organizations are expected to make accurate decisions regarding their market, operations and investment. This points to having evidence-based decision-making approaches. Big data analytics, or BDA, has been embraced by various business sectors worldwide as an enabler for such evidence-based decision making. Big data or BD is defined as a collection of structured, unstructured and semi-structured data gathered from different sources and applications.

The insurance world is looking at a tech transformation by adopting new technologies which are expected to enable the insurers to drastically improve risk assessment, personalize policies, streamline claims processing, and deliver more customer-centric services that stand out in the ever-evolving digital landscape by leveraging real-time data from connected devices, leading to a more efficient and data-driven approach to insurance operations. In order to have a strong data analysis and management structure that will guarantee better customer service, reduce expenses, and have more robust information available not only for financial reporting but also to have insights on business trends and products, customer expectations and investment options, an increasing number of insurers are now opting to invest in modern technology.

Big data analytics has therefore become the most powerful buzzword in almost all the business sectors. [1] defines big data analytics as techniques that are deployed to uncover hidden patterns and bring insight into interesting relations in understanding contexts by examining, processing, discovering, and exhibiting the result.

Larger new data sources are being generated from mobile devices and large companies such as Facebook, Apple, Google, Yahoo, X who are beginning to look carefully at this data to find useful patterns to improve user experience. It is undeniable that millions and millions of data sets are now being generated from technology devices and, as a growing field, many digital apps are also contributing to this phase of data collection. The real success of any industry can be measured based on how the big data is analysed, potential knowledge is discovered, and productive business decisions are made.

Big data analytics said to be more valuable to those companies, focusing on getting insight into customer behaviour, trends and patterns. Modern companies therefore need new approaches on how to analyse different big data which is unstructured data by incorporating new data analysis approaches such as artificial intelligence (AI), association rule learning, machine learning, genetic algorithm, classification tree analysis, social network analysis, regression analysis, and

sentiment analysis. According to [2], the performance of big data in the insurance market results in cost reduction, better access to insurance services, and more fraud detection that benefits the customers and stakeholders.

Big data is at the heart of the insurance industry through the uses it provides, and it is considered as the beginning of the "digital revolution" when humans have able to store more digital information in technological tools than ever before [3]. Big Data analytics for insurance would involve the systematic use of data and advanced analysis techniques to gain insights into customer behaviour, risk assessment, claims patterns, and market trends, enabling insurers to make informed decisions regarding product development, pricing, underwriting, fraud detection, and customer service, ultimately aiming to optimize operations and improve profitability. Big data analytics is therefore expected to revolutionize the insurance industry not only through improved customer experience, risk assessment, operations, and fraud detection but also as a critical tool to improve decision-making and financial performance. Big Data analytics is evolving rapidly, the various types of data being processed are huge, and ensuring that this data is being used efficiently is becoming increasingly more difficult.

2. Big Data and Insurance Industry in Kenya

Although various companies in different sectors in Kenya are still exploring big data technology, it is not yet widely adopted across the business landscape. Some of the companies that have adopted big data analytics are in the telecommunications sector where it is playing a crucial role among other things optimizing network performance, product development, enhancing customer experience, identifying fraudulent activity, and developing targeted marketing strategies [4]. These has enabled telcos to gain a competitive edge by analysing vast amounts of customer data to improve service quality, customer retention, and overall operational efficiency; with prominent examples like Safaricom leveraging data to personalize offers and optimize network coverage across different regions in Kenya [4, 5].

The insurance industry in Kenya on the other side plays a significant role in the growth and development of the financial services. Although it is identified in the Vision 2030 as critical in the transformation of Kenya under the economic pillar [6], the insurance sector in Kenya is riddled by many challenges among them low penetration levels, low persistency ratios and poor claims payment history [7]. These problems have been persistent for a long time, and they are mainly attributed to very low innovation levels across the sector.

Despite its role in financial inclusion and sustainable economic growth, the insurance penetration in Kenya has been on the decline over recent years. As of 2021, the insurance penetration rate in Kenya stood at 2.33 per cent in 2022 with a marginal increase from 2.29 in 2021 [8] which is significantly

lower than the global average of around 7%. The relatively lower level of insurance penetration in Kenya has been attributed to several factors including traditional insurance products, perceived low rate of returns for life policies and cumbersome claim settlement procedures among many factors. Other factors include lack of consumer knowledge and awareness, negative perceptions, cultural and religious beliefs, inappropriate products, and limited distribution channels [9]. To address these, over the years insurance companies have been trying on how best they can offer individualized services to their clients and products that combine protection against risk with opportunities for financial gain.

New and innovative insurance start-ups with digital-first mindset are cropping up with digital space being their primary target. As a result, the industry has experienced financial reforms, technological advancements, and globalization to improve its efficiency, productivity, market structure, and performance. Legacy insurance providers are under pressure to invest in digital markets and acquire new expertise to keep up with the rapidly changing markets and population demographics. Increased visibility, reduced sales costs and increased penetration rates are some of the benefits incentivizing the rapid uptake of digitalization and technology.

Insurance consumers today are glued up on technology and prefer to have their world revolve around their smart phones, insurers have seen opportunities that exist especially in engaging their customers and prospects through social media and digital platforms. This has led to insurers finding themselves handling huge amounts of data which is in different formats and which the existing data mining techniques cannot be applied. Therefore, more advanced techniques or methods are required to manage, process, and generate valuable information from this massive amount of data. The formation of this massive amount of data through divergent sources has led to high volume, high velocity, and a variety of data that has led to the coining of the term big data.

Given that insurance business is majorly based on the analysis of data to understand and effectively evaluate risk. Actuaries and underwriting professionals depend upon the analysis of data to be able to perform their core roles. With the advent of various web technologies, mobile devices, and sensor devices, the insurance sector is experiencing continuous growth in data volume. The unprecedented increase in data through internal and external sources has resulted in a massive amount of data. This high volume of data brings challenges to the data itself such as the storage of the data for processing is not possible through traditional tools and thus more innovative methods should be developed to handle this data deluge. Furthermore, a smooth data analytics process requires a meticulous approach to developing disruptive solutions that customers highly anticipate.

Big data has therefore inspired insurance companies in Kenya as a new tool to help insurance companies improve competitiveness of their products and services through a systematic use of data and advanced analysis techniques to

gain insights into customer behavior, risk assessment, claims patterns, and market trends, enabling insurers to make informed decisions regarding product development, pricing, underwriting, fraud detection, and customer service, ultimately targeting to optimize operations and improve profitability. Big Data Analytics in the insurance industry is becoming a potential field for giving insights from extremely huge data sets, as well as improving results while simultaneously cutting costs [10].

The data-driven revolution is not merely boosting efficiency but fundamentally reshaping how insurers understand and serve their clients, offering more personalized and responsive services than ever before. Although big data analytics in insurance industry across the globe has gradually become a cornerstone for driving transformation, innovation, and efficiency and reshaping operational performance, basic research is still lacking in academic publications on its adoption, use and impact in the insurance industry in Kenya.

3. Characteristics of Big Data

Big Data has been differentiated into several characteristics and many researchers have been developing more characteristics for new purposes over the past years. Although there is no unanimous agreement on the characteristics of big data, the term is coined mainly to refer to data that is characterized by being generated at high speed and in large volume, coming from a wide variety of sources that provide veracity about the information and provide value for decision making.

Initially, the differentiation involved three main characteristics that included Volume, Velocity and Variety [11]. While many researchers believed in three and four Vs for BD characteristics, [12] reviewed the big data characteristics and added Value to the existing four Vs (Volume, Velocity, Variety and Veracity) making it five Vs. [13] added Validity and Volatility to the existing characteristics making the Vs seven in number. The seven Vs included Volume, Velocity, Variety, Veracity, Value, Validity, and Volatility. It is [14] who raised the big data characteristics to ten by adding Variability, Visualization and Vulnerability to this list of existing characteristics from previous studies making them ten Vs of BD. The 10 Vs are: Volume, Velocity, Variety, Veracity, Value, Validity, Volatility, Variability, Visualization, and Vulnerability. Big data analytics has some striking characteristics that relate to insurance.

Volume

As digital engagement continues rising worldwide, with users embracing online shopping, streaming videos and other content, and social media, so does the amount of data humans generate daily. According to [15], there were 5.56 billion internet users worldwide, which amounted to 67.9 percent of the global population as of February 2025. Of this total, 5.24 billion, or 63.9 percent of the world's population, were social media users. All those users are expected to drive significant data growth each year. If the volume of data is very large, then

it is considered as a 'Big Data'. Insurers deal with vast amounts of data, including historical claims data, policyholder information, and external data sources like weather patterns and social media trends. This means whether a particular data can be considered as Big Data or not, is dependent upon the volume of data. Therefore, while dealing with Big Data it is necessary to consider a characteristic 'Volume'.

Variety

Not only does the amount of data generated daily characterize big data, but its variety is much larger. Nearly every facet of digital life generates data in some form. From social media posts and videos to consumer data and medical records, almost every aspect of life is now tracked and stored. Data also comes in three ways: structured, semi-structured, and unstructured data sets.

Structured data: This data is basically organized data. It generally refers to data that has defined the length and format of data. structured data can be crudely defined as the data that resides in a fixed field within a record. Structured data is split into multiple tables to enhance the integrity of the data by creating a single record to depict an entity. The business value of structured data lies within how well an organization can utilize its existing systems and processes for analysis purposes.

Semi- Structured data: This data is basically semi-organized data. It is generally a form of data that does not conform to the formal structure of data. Log files are examples of this type of data. Semi-structured data is not bound by any rigid schema for data storage and handling. The data is not in the relational format and is not neatly organized into rows and columns like that in a spreadsheet. Since semi-structured data doesn't need a structured query language, it is commonly called *NoSQL data*. Insurers leverage data from various sources, including internal systems, external databases, and IoT devices.

Velocity

The speed of data generation is another distinctive characteristic of big data, which is influenced by the volume and variety of information. Thanks to smartphones and social media, people are producing data at an unprecedented pace, especially since the emergence of big data systems. Velocity describes the rapid accumulation of data, characterized by a significant and ongoing influx. This aspect of big data indicates how quickly data is created and processed to satisfy various demands. Insurers need to process and analyse data in real-time to make informed decisions and respond quickly to events. Big data helps automate and speed up claims processing, leading to faster payouts and improved customer satisfaction.

Veracity

Veracity, one of the key characteristics of big data, focuses on the trustworthiness and reliability of the data. Veracity refers to the quality, accuracy, and reliability of collected data, ensuring the data is trustworthy and useful for analysis and decision-making. It addresses issues like noise, inconsisten-

cies, and errors that can undermine the value of big data. Accurate and reliable data is crucial for making informed decisions and drawing valid conclusions from big data analysis. Data validation, cleaning, and integration are some of the strategies that can be used in improving veracity of data. The accuracy and reliability of data are crucial for making sound decisions in the insurance industry. Insurers need to implement robust data cleaning and validation processes to ensure data quality.

Value

An important characteristic of big data in this industry is value, how can a business not only collect and manage big data, but how can the data which holds value be identified and how can organizations forward-engineer (as opposed to retrospectively evaluate as is normally the case for insurance industry) commercial value from the data. The bulk of Data having no Value is of no good to the company, unless you turn it into something useful. Data in itself is of no use or importance, but it needs to be converted into something valuable to extract information. Value therefore describes the benefits the data provides to an organization, including improved decision-making, enhanced customer experience, increased operational efficiency, and the potential for new revenue streams. Big data can help identify inefficiencies and optimize processes thus reducing costs across various departments in the organization.

Big data analytics provide insurers with valuable insights that enable them to make data-driven decisions. By analysing vast amounts of data, insurers can assess risks more accurately and set premiums accordingly. It also enables insurers to offer personalized products and services that better meet the needs of individual customers by understanding customer behavior and preferences through big data analytics.

Visualization

In the era of big data, visualization is a crucial tool for extracting meaningful insights from vast datasets and making informed decisions. Given the volume, variety, and velocity of big data, traditional visualization methods often fall short, requiring more sophisticated approaches to make sense of such vast amounts of information. Visualization uses visual elements like charts, graphs, and maps to represent complex data, making it easier to understand and interpret trends, patterns, and insights that might otherwise be hidden in raw data. Visualization is the creation of visual guides or references so businesses, analysts, and other users can "see" their data in a digestible format. Interactive visualizations can engage users more effectively compared to static reports. Dashboards and real-time visual data representations allow users to interact with the data, drilling down into specifics and exploring different scenarios, leading to deeper insights and more meaningful engagement with the information. Visualizing in big data analytics can help in identifying patterns, correlations, and outliers that can help business to understand the underlying trends and relationships in various business aspects. By choosing appropriate techniques and tools, insurance compa-

nies can enhance their ability to analyse, understand, and act on complex datasets in real-time.

Variability

Variability is often considered one of the key characteristics of big data. Variability refers to the inconsistencies and unpredictable changes in data, including fluctuations in data flow, changing data formats, or shifts in how data is interpreted over time, which can make it challenging to maintain accurate and reliable models. Data scientists therefore need to design models that can adapt to changing conditions, ensuring their systems are resilient.

4. Big Data Use in Insurance Industry

Big data has contributed to the development of the insurance industry in recent years and has become an essential part of the various traditional and new processes of this industry [3]. Today, insurers across various sectors such as travel, health, life, and property are leveraging Big Data to revolutionize their operations. This data-driven revolution is not merely boosting efficiency but fundamentally reshaping how insurers understand and serve their clients, offering more personalized and responsive services than ever before. Key applications of data analytics in insurance:

Risk assessment:

At the core of insurance lies risk assessment process which big data analytics is revolutionizing through predictive analytics. Analysing customer data (demographics, driving records, credit history) to accurately assess risk and determine appropriate premiums for insurance policies. Insurers can use data analytics to analyse risk, enabling them to more efficiently price insurance and reduce losses. AI (artificial intelligence) and ML (machine learning) algorithms sift through these datasets to predict claims, detect fraud, and identify trends [16]. Accurate risk prediction helps insurers better price policies and mitigate potential losses, benefiting both the insurer and the insured. By leveraging predictive analytics, insurance firms excel in risk management, balancing risk assessment and customer validation. For example, an insurance may analyse the likelihood of damage from floods or other natural catastrophes and modify rates as necessary using data on weather patterns and past claims data. Better risk assessment and improved tools for policyholder engagement or interaction can enhance the effectiveness of risk reduction advice and provide new mechanisms to deliver risk mitigation advice and services to policyholders [17, 18].

Claims processing:

Processing claims has been historically a repetitive, time-consuming task, that involves the thorough analysis of every specific case and the information provided by the insured. This process often leads to human error and serious bottlenecks, and the likelihood of suffering fraud is relatively high. Traditional systems face a major disadvantage because they cannot handle rising claim complexity as well as increased claim volume especially within health and life in-

surance [19]. Insurance companies therefore spend a lot of money on insurance claims processes.

Big data analytics can be utilized to speed up and increase the precision of claim processing. The application of machine learning (ML) technology serves as an innovative solution which brings automated claim triage to insurance companies for greatly improved and expedited decision-making effectiveness [20]. Insurance companies may lower fraud by examining fraud claim trends and put more control mechanisms to prevent them. AI is enabling insurance companies to apply machine learning algorithms to the legacy processes to improve claims handling and enhance customer experience. Advanced algorithms of AI handle the initial claims routing, thereby increasing its accuracy and efficiency. Using AI, an insurance company can decrease its hiring expenses by automating time-consuming payouts and claim management. Therefore, the time needed to process these claims is reduced from days to hours or even minutes, thereby providing customized contracts for the customers.

Fraud Detection:

Insurance fraud can be costly for insurance companies, as they may end up paying out claims that they would not have approved if they had accurate information. This, in turn, can lead to higher premiums for policyholders as insurance companies seek to recoup their losses. Insurance fraud is the act of intentionally deceiving an insurance company to receive an illegitimate benefit. There are various forms of insurance fraud which include claims and underwriting insurance fraud. Claims insurance fraud involves deceiving an insurance company to receive an illegitimate benefit. In contrast to underwriting insurance fraud, which involves providing false information during the insurance application process, claims insurance fraud occur after a policy has been issued and a claim is submitted. One of the main advantages of big data is that it allows for the creation of detailed profiles of individuals and entities. By collecting and analyzing data from a wide range of sources, insurance companies can gain a comprehensive understanding of a person's behaviour and history, making it easier to identify fraudulent activity. Big data algorithms can identify unusual claim patterns, such as sudden increases in the frequency or type of claims filed by a specific individual or group [21]. By analyzing historical and real-time data, insurers can also identify potential fraud early in the claims process, allowing for timely intervention and investigation. [17, 18].

Customer segmentation and personalization:

Grouping customers based on their characteristics to deliver targeted marketing campaigns and customized insurance products. By employing advanced analytics, businesses can segment customers more precisely, offering tailored marketing strategies that cater to specific needs and preferences [22]. Techniques such as data mining, machine learning, and predictive analytics provide robust frameworks for understanding consumer behaviour patterns and preferences, thus driving personalized marketing strategies [23]. Machine learning

models can be used to predict the uptake of insurance and therefore increase insurance penetration which is very low in the Kenyan market. Insurance companies need to understand how the advancement in robotics will change the customer experience and enable new products. For example, chatbots can offer 24-hour assistance and provide a personalized experience where humans can fail. According to [24], big data serves as a cornerstone for predictive customer segmentation by utilizing advanced analytics to identify patterns and forecast consumer behaviours more precisely than traditional methods. By leveraging Big Data, insurers can gain deep insights into customer needs, allowing them to craft plans that are precisely tailored to those requirements.

Product development:

One-size-fits-all insurance policies are becoming a thing of the past. Big data empowers insurers to tailor policies to individual needs. By analysing customer data and behaviour, insurers can create personalized offerings that better match specific lifestyles, preferences, and risk profiles. Insurers can employ data analytics to better understand their customers' requirements and preferences and therefore create insurance products/packages that uniquely suit each customer's wishes and preferences [18]. For instance, an insurer may utilize information on a client's driving patterns to create a personalized auto insurance plan that charges less for careful drivers. This leads to enhanced customer experience through personalized products and services based on individual needs.

Pricing and Underwriting:

The underwriting process aims to set a fair premium based on individuals' risk profile. This has been traditionally a tedious and time-consuming process, that requires the collection and evaluation of vast amounts of data, often sorted chaotically and fragmentally. Traditional underwriting is prone to errors and complexities, making the process lengthy, inefficient and cumbersome. Limited capacity to address underwriting often leads to higher premiums and a lack of personalization. Utilizing data to dynamically adjust premiums based on real-time risk factors and market conditions. One of the most profitable ways where BDA can upgrade the insurance industry comes under underwriting and pricing. By incorporating data from multiple sources, insurers can better evaluate applicants and make more informed decisions, resulting in lower risk and higher profitability. This vast data enables the creation of pricing models that not only fit the client's budget but also ensure the company's profitability. It also enables insurers to make reliable decisions tailored to the buyer's risk and coverage needs. As a result, insurance companies can now create pricing models for pricing their policies more accurately and competitively for each customer. AI in underwriting minimizes the possibility of human error by combining massive datasets in various formats, making them less prone to mistakes [25]. AI is algorithmically bound to be self-reliant and learn from previous mistakes, saving time and making the process more efficient and scalable. It can also help insurance companies to get information about a person's

location, age, daily routine, medical history, and the likelihood of filing a claim. Insurance companies can use this data to set premiums that suit their needs. To successfully modernize the underwriting and customer onboarding process, businesses must be restructured, and agile principles must be applied.

Regulatory compliance:

As organizations increasingly leverage Big Data technologies, the importance of regulatory compliance and ethical considerations is becoming paramount [26]. Compliance with existing regulations and adherence to ethical principles are critical for ensuring responsible and transparent use of big data technologies. Big Data is revolutionizing how insurers stay compliant with ever-changing regulations. Through advanced analytics and real-time data processing, insurers can ensure precise reporting, continuous monitoring, and proactive issue resolution. This minimizes the risk of penalties and boosts transparency. Moreover, Big Data automates compliance workflows, cutting down on administrative tasks and streamlining regulatory adherence. By strategically leveraging big data, insurers can confidently navigate the complex regulatory landscape with ease.

Customer Retention:

No business likes to lose its customer base. A business is considered successful if its customer retention rate is higher. The insurance industry is no exception. So, it utilizes big data to identify potential clientele and proactively take measures to retain them. Based on customer activity, algorithms can also predict the early signs of client displeasure. Using the information provided, companies can quickly react to improving customer experience and be able to address customers' complaints. Insurers can offer discounts or even change the pricing model for the client. The integration of telematics data, natural language processing, and other innovative technologies into underwriting workflows offers new opportunities for insurers to enhance customer experience and drive competitive advantage in the dynamic insurance marketplace [27]. Building customer loyalty is important because it helps the business sector to improve the brand image as the brand has already earned the trust of loyal customers, and they are more likely to share a positive experience than a new customer [3].

Cost Reductions:

The combination of human mistakes with inconsistent choices and unmodernized procedures leads to slower operations backed by elevated operational costs [28]. Cost-cutting is one of the many benefits of leveraging technology. The increased role of machines in the industry increases efficiency which eventually leads to cost reductions. The cost-effectiveness underscores the transformative potential of AI in revolutionizing database management practices. According to [29], the transformative potential of AI not only enhances the user experience by reducing latency but also optimizes resource allocation, resulting in cost savings for organizations. Big data technology can be leveraged to automate manual processes, making them more efficient and

reducing the costs spent on handling claims and administration. This will allow the companies to offer lower premiums to their clients and hence stand tall in the competitive market. The implementation of this technology provides organizations with important savings opportunities. The elimination of human involvement in claim processing leads insurance companies to reduce their operational costs and improve customer satisfaction through rapid claim resolution [30].

Operation Efficiency:

Big data enhances operational efficiency by optimizing internal processes. Insurers can identify bottlenecks, streamline workflows, and improve overall productivity, leading to cost savings and better resource management. Using Big Data, Machine Learning algorithms, and predictive modelling, insurers can develop models that identify inefficiencies and areas for improvement. These models can help insurers automate certain processes, reduce manual intervention, and identify potential issues before they arise. Using data analytics on various aspects of their operations, insurers can make data-driven decisions that improve their overall efficiency and reduce costs [16]. Additionally, by automating certain processes and reducing manual intervention, insurers can process claims more quickly, resulting in improved customer satisfaction and loyalty. Data management using Big Data Analytics propels improved intelligent organizational effectiveness and decision-making [31]. The synergy between Big Data and AI has the potential to unlock valuable insights, improve decision-making processes, and enhance overall efficiency in various industries [18, 26].

5. Big Data Technologies

According to [32], approximately 95% data is present in the form of raw data (not in any form) which brings many challenges to businesses and enterprises. To make big data analysis process more accurate, fast, and precise several innovative tools and techniques have been incorporated and practiced. Insurance companies have embarked on a journey of profound changes, where investment in new technologies is key to driving profits and surviving in an ever-evolving competitive landscape. Big Data Technology refers to the hardware and software tools that are used to manage types of datasets and transform them into useful data for businesses. This technology analyses, processes, and extracts valuable information from a huge set of data containing complex structures. Big data technology is widely connected with emerging and latest technologies like cloud computing, Machine Learning (ML), Artificial Intelligence (AI), and the Internet of Things (IoT). These tools and techniques discover the hidden pattern, find the unknown-correlation, and extract meaningful information from the data.

1) Cloud Computing

Cloud computing is the on-demand delivery of computing resources, like servers, storage, and applications, over the internet, allowing users to access and use them without

needing to own or manage the underlying infrastructure. It provides scalable storage and computing resources, enabling organizations to handle large datasets and complex analyses without the limitations of traditional on-premise infrastructure. Using different services, cloud computing offers more benefits than traditional computing. Cost saving, scalability, mobile storage, anywhere access, better security, energy saving, environment benefits are some of benefits of cloud computing. It also offers increased flexibility, and faster deployment of big data solutions. Clients typically pay only for cloud services that they use, helping them lower their operating costs, run their infrastructure more efficiently, and scale as their business needs change. Cloud-based platforms come with built-in redundancy that can save the business from loss of the data. Loss of data leads to loss of productivity, revenue, and brand reputation [33]. Cloud computing also gives high security by utilizing information encryption and solid access controls [34]. Cloud computing enables insurers to analyse large datasets in real-time and with greater depth, leading to faster and more accurate predictions. Cloud analytics can help predict customer financial moves and needs, allowing insurers to proactively reach out and offer alternative products. Cloud computing offers cost savings, scalability, accessibility, and enhanced collaboration, making it a popular choice for insurance businesses.

2) Machine learning

Machine learning (ML) is the science of computers running without being explicitly programmed. It uses a series of statistical techniques, such as mathematical modelling, data visualization, and pattern recognition, to conduct self-learning activities with input data to predict and understand data trends and patterns [35]. Machine learning plays a powerful role in data analytics by automating tasks, uncovering hidden patterns, and making predictions from large and disparate data sets. ML algorithms can identify patterns, predict outcomes, and automate tasks in data analysis, making it a powerful tool for extracting insights from big data. Once machine learning models are trained on historical company data, they can predict future outcomes to help minimize customer churn, build targeted marketing campaigns, and set optimal pricing levels. Machine learning has predictive algorithms by data interpretation, which is followed by learning algorithms in an unstructured program. According [36], there are three main categories of ML which include supervised, unsupervised, and reinforcement learning, which is done during “data preprocessing,” “learning,” and “evaluation phase.” According to the researcher preprocessing is related to transformation of raw data into the right form that can be deployed in the learning phase, which comprises of some levels like cleaning the data, extracting, transforming, and combining it. In the evaluation phase, data sets will be selected, and evaluation of performance, statistical tests, and estimation of errors or deviation occur. The current insurance claims processing evolution demands the use of machine learning (ML) in anomaly detection as a way to enhance operational efficiency [20].

Insurance underwriting accuracy has also been increasingly influenced by advancements in data analytics and machine learning techniques. Insurers utilize vast amounts of data, including historical claims data, socio-economic indicators, and health records, to develop sophisticated underwriting models that accurately predict risk. Machine learning methods encompass a diverse set of algorithms and techniques that enable computers to learn from data and make predictions or decisions without explicit programming. Decision trees can be employed to analyse applicant information and assess risk factors, aiding underwriters in making more informed decisions. In insurance underwriting, logistic regression, which is a type of ML can help evaluate the likelihood of an applicant filing a claim or defaulting on premiums.

3) Artificial Intelligence (AI):

Artificial Intelligence (AI) refers to the ability of machines to perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. AI encompasses various technologies and techniques that enable computers to process data, recognize patterns, and make predictions. AI is the outcome of successful applications of big data and ML technologies to understand the past and predict the future based on enormous data [35]. Like other financial service industries, the insurance sector is one of the most data-rich sectors. This data helps carriers decide what insurance to give to which people and which premiums they should charge. Artificial intelligence can improve providers' decision-making capabilities, driving increased care to their customers while improving their bottom lines. Artificial Intelligence can be used in enhancing underwriting and claims processes. AI can now analyse claimants' data, and make personalized recommendations based on that information. This technology can improve the accuracy of risk assessments and help with the correct payout for that claimant. It also helps to flag fraudulent claims and accelerate the claims process. Another form of AI applicable to the insurance industry is Robotic process automation (RPA). RPA is a repetitive and automated process which can be used to imitate human behaviour and carry out repetitive tasks such as sending e-mails, complete spreadsheets, and record and re-enter data for other tasks [37].

4) Blockchain:

Blockchain technology is an advanced database mechanism that allows transparent information sharing within a business network. In this technology, transactions are grouped into "blocks" and each block is linked to the previous block, creating a chain. The technology has its roots in the late 1980s when a computer scientist named Ralph Merkle patented Hash trees or Merkle trees [38]. These trees are a computer science structure for storing data by linking blocks using cryptography. Blockchain is a disruptive ledger technology that has sparked a significant interest in supporting big data systems with high security and efficient network management [39]. Technology is normally decentralized and distributed in nature where instead of a single entity controlling a transac-

tion, it gets maintained across a network of computers, ensuring no single point of failure or control. The data is therefore chronologically consistent because you cannot delete or modify the chain without consensus from the network. The system has built-in mechanisms that create an audit trail and prevent unauthorized transaction entries and create consistency in the shared view of these transactions [40].

Blockchain with its decentralization and security nature has the great potential to improve big data services and applications [39]. The blockchain reduces the dangers of data being kept centrally by storing it across its network. Its network is devoid of centralized sources of vulnerability that may be exploited by computer hackers [41]. Technology therefore provides a secure and transparent platform for storing and sharing insurance data, reducing the risk of fraud and errors. Through blockchain technology various insurance processes such as policy issuance and claim payments can be automated leading to greater efficiency and reduced costs. As insurance companies discover and implement new applications, blockchain technology continues to evolve and grow. Companies are able handle limitations of scale and computation, and potential opportunities are limitless in the ongoing blockchain revolution.

5) Internet of Things (IoT):

There is no general definition approved by the majority or by the global user community, and therefore the Internet of Things is maturing and continuing to be the newest, most popular concept in the world of information technology. According to [42], Internet of Things (IoT) describes a vast array of objects with sensing and actuating devices that collect, analyse and share data across other objects, programs, and platforms. IoT is created to enable information and data exchange between objects which are usually embedded with sensors technology to capture data from their surroundings [43]. describes it as a network of physical objects that are digitally connected to sense, monitor, and interact within a company and between the company and its supply chain enabling agility, visibility, tracking, and information sharing to facilitate timely planning, control, and coordination of the supply chain processes.

The IOT technology has made it possible to collect, for example, customer data using a mobile phone. Today smartphones record data, voice, video, audio, motion, location and much more. Smart watches capture steps taken, burned calories, heart rate, temperature, stress level, sleeping habit and other personal data. They display real weather conditions, news etc. Moreover, the devices can be used to pay for goods, hold air tickets and boarding passes and much more. It is therefore now possible for insurance companies to obtain real-time customer's data and analyse it by utilizing IoT and big data. This enables insurers to enjoy several benefits such pricing of products and premium calculation in a more accurate and fairer for each customer based on their risk. By adjusting premiums in real-time based on changing risk profiles and market conditions, insurers could achieve better risk

management outcomes.

6. Summary and Conclusions

Advancements in technology, including the Internet of Things (IoT) and further sophistication of AI and machine learning, will deepen the reliance on data-driven insights and decisions in various sectors. Additionally, collaborations with tech innovators and the integration of emerging data sources will open new avenues for the various sectors of the economy to innovate their offerings and services. Big Data Analytics applications have become indispensable tools for organisations seeking a competitive edge through data-driven decision-making.

The insurance industry's future will be shaped by technological advancements, focusing on personalized products, AI-powered claims processing, and data analytics to improve customer experience and risk management. With the ability to analyse large volumes of data, insurance organizations can make data-driven decisions that improve strategy, enhance customer experiences, and optimize resources.

While there is a growing body of literature on the theoretical aspects of big data and its applications in various domains, there is limited research that focused on evaluating its adoption and effect in the insurance industry. Future research should therefore focus on taking an assessment on the big data adoption and its impact on the insurance industry while addressing implementation challenges.

Abbreviations

AI	Artificial Intelligence
AKI	Insurance Association of Kenya
ML	Machine Learning
BD	Big Data
BDA	Big Data Analytics
IoT	Internet of Things

Author Contributions

Cyprian Omenge Nyambane is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

- [1] Wang, H., Cui, Z., Sun, H., Rahnamayan, S., & Yang, X. S. (2017). Randomly attracted firefly algorithm with neighborhood search and dynamic parameter adjustment mechanism. *Soft Computing*, 21, 5325-5339.
- [2] Rana, A., Bansal, R., & Gupta, M. (2022). Big data: A disruptive innovation in the insurance sector. In *Big data analytics in the insurance market* (pp. 165-183). Emerald Publishing Limited.
- [3] Belhadi, A., Abdellah, N., & Nezai, A. (2023). The effect of big data on the development of the insurance industry.
- [4] Kariuki, J. T., & Kagiri, A. (2018). Strategic role of big data analytics on innovation in the telecommunications sector in Kenya: A Case of Safaricom PLC. *The Strategic Journal of Business and Change Management*, 5 (4), 1573-1592.
- [5] Kogi, M. N., (2023). The Role of Big Data Analytics in Enhancing Customer Relations in A Telecommunication Company in Kenya: A Case of Safaricom: Daystar University, School of Communication.
- [6] GOK. (2012). Sessional paper Number 10 of 2012 On Kenya Vision 2030, Office of the Prime Minister Ministry of state for Planning, National Development and Vision 2030. *Government Press*.
- [7] Ipomai, R. (2016). *Adoption of Business Intelligence Solutions: a Case of Kenyan Insurance Industry* (Doctoral dissertation, University of Nairobi).
- [8] AKI. (2022) Insurance Industry Market Report 2022. *Association of Kenya Insurers*.
- [9] Saeed, M., & Arshed, N. (2023). Big Data Analytics Technology Adoption in Kenya Insurance Industry: A Systematic Literature Review. *Asian Bull. Big Data Manag*, 3, 1-17.
- [10] Banu, A. (2022). Big data analytics—tools and techniques—application in the insurance sector. In *Big data: A game changer for insurance industry* (pp. 191-212). Emerald Publishing Limited.
- [11] Fan, W., & Bifet, A. (2013). Mining big data: current status and forecast to the future. *ACM SIGKDD explorations newsletter*, 14(2), 1-5.
- [12] Nguyen, T. (2018) "A Framework for Five Big V's of Big Data and Organizational Culture in Firms", 2018 IEEE International Conference on Big Data (Big Data), 2018. Available: <https://doi.org/10.1109/bigdata.2018.8622377>
- [13] Uddin, M. F., & Gupta, N. (2014, April). Seven V's of Big Data understanding Big Data to extract value. In *Proceedings of the 2014 zone 1 conference of the American Society for Engineering Education* (pp. 1-5). IEEE.
- [14] Ranjan, J. (2019) "The 10 Vs of Big Data framework in the Context of 5 Industry Verticals", *PRODUCTIVITY*, vol. 59, no. 4, pp. 324-342, 2019.
- [15] Statista (2025). *Worldwide digital population 2025 report*. Retrieved March 17, 2025, from <https://www.statista.com/statistics/617136/digital-population-worldwide/>
- [16] Ho, C. W., Ali, J., & Caals, K. (2020). Ensuring trustworthy use of artificial intelligence and big data analytics in health insurance. *Bulletin of the World Health Organization*, 98(4), 263.

- [17] Dahiya, M., Sharma, S., & Grima, S. (2022). Big data analytics application in the Indian insurance sector. In *Big data analytics in the insurance market* (pp. 145-164). Emerald Publishing Limited.
- [18] Asthana, N., & Osama, M. (2024). Data Analytics in Insurance Product Management. *Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006-4023*, 6(1), 594-599.
- [19] Carvalho, D. V., Pereira, E. M., & Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. *Electronics*, 8(8), 832.
- [20] Malali, N. (2022). Using Machine Learning to Optimize Life Insurance Claim Triage Processes Via Anomaly Detection in Databricks: Prioritizing High-Risk Claims for Human Review. *International Journal of Engineering Technology Research & Management (IJETRM)*, 6(06).
- [21] Danenas, P., & Garsva, G. (2014). Intelligent techniques and systems in credit risk analysis and forecasting: a review of patents. *Recent Patents on Computer Science*, 7(1), 12-23.
- [22] Poonia, S., Kumar, A., & Shrivastava, S. (2025). Predictive Customer Segmentation and Personalization through Big Data in Digital Marketing. *Innovate to Dominate: AI and Sustainability in Business*, 117.
- [23] Osakwe, J., Shilongo, A., & Ziezo, M. (2023). Optimising Customer Segmentation in Digital Marketing Using Predictive Analytics: A Review of Literature. Available at SSRN 4662191.
- [24] Asif, S. (2025). Leveraging Big Data and Predictive Analytics in Marketing Strategies. *Innovate to Dominate: AI and Sustainability in Business*, 29.
- [25] Javanmardian, K., Ramezani, S., Srivastava, A., & Talischi, C. (2021). How data and analytics are redefining excellence in P&C underwriting. *McKinsey & Company*, Sep, 24.
- [26] Blessing, E. (2024). Regulatory Compliance and Ethical Considerations: Compliance challenges and opportunities with the integration of Big Data and AI.
- [27] Bishop, N. (2024). Application of Machine Learning Techniques in Insurance Underwriting. *Journal of Actuarial Research*, 2(1), 1-13.
- [28] Leo, M., Sharma, S., & Maddulety, K. (2019). Machine learning in banking risk management: A literature review. *Risks*, 7(1), 29.
- [29] Gadde, H. (2023). Leveraging AI for Scalable Query Processing in Big Data Environments. *International Journal of Advanced Engineering Technologies and Innovations*, 1(02), 435-465.
- [30] Asri, H., Mousannif, H., & Al Moatassime, H. (2019). Reality mining and predictive analytics for building smart applications. *Journal of Big Data*, 6, 1-25.
- [31] Niu, Y., Ying, L., Yang, J., Bao, M., & Sivaparthipan, C. B. (2021). Organizational business intelligence and decision making using big data analytics. *Information Processing & Management*, 58(6), 102725.
- [32] Rawat, R., & Yadav, R. (2021). Big data: Big data analysis, issues and challenges and technologies. In *IOP Conference Series: Materials Science and Engineering* (Vol. 1022, No. 1, p. 012014). IOP Publishing.
- [33] Islam, N. (2017). Review on Benefits and Security Challenges of Cloud Computing. *International Journal of Computer Science and Information Technologies*, 8(2), 224-228.
- [34] Aldossary, S., Allen, W. (2016). Data security, privacy, availability and integrity in cloud computing: issues and current solutions. *International Journal of Advanced Computer Science and Applications*, 7(4), 485-498.
- [35] Mutlag, A. A., Abd Ghani, M. K., Arunkumar, N. A., Mohammed, M. A., & Mohd, O. (2019). Enabling technologies for fog computing in healthcare IoT systems. *Future generation computer systems*, 90, 62-78.
- [36] Yu, H., Yang, X., Zheng, S., & Sun, C. (2018). Active learning from imbalanced data: A solution of online weighted extreme learning machine. *IEEE transactions on neural networks and learning systems*, 30(4), 1088-1103.
- [37] Enríquez, J. G., Jiménez-Ramírez, A., Domínguez-Mayo, F. J., & García-García, J. A. (2020). Robotic process automation: a scientific and industrial systematic mapping study. *IEEE Access*, 8, 39113-39129.
- [38] Merkle, R. C. (1987, August). A digital signature based on a conventional encryption function. In *Conference on the theory and application of cryptographic techniques* (pp. 369-378). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [39] Deepa, N., Pham, Q. V., Nguyen, D. C., Bhattacharya, S., Prabadevi, B., Gadekallu, T. R., & Pathirana, P. N. (2022). A survey on blockchain for big data: Approaches, opportunities, and future directions. *Future Generation Computer Systems*, 131, 209-226.
- [40] Muheidat, F., Patel, D., Tammisetty, S., Tawalbeh, L. A. A., & Tawalbeh, M. (2022). Emerging concepts using blockchain and big data. *Procedia Computer Science*, 198, 15-22.
- [41] Chattu, V. K. (2021). A review of artificial intelligence, big data, and blockchain technology applications in medicine and global health. *Big Data and Cognitive Computing*, 5(3), 41.
- [42] Koohang, A., Sargent, C. S., Nord, J. H., & Paliszkievicz, J. (2022). Internet of Things (IoT): From awareness to continued use. *International Journal of Information Management*, 62, 102442.
- [43] Ben-Daya, M., Hassini, E., & Bahrour, Z. (2019). Internet of things and supply chain management: a literature review. *International journal of production research*, 57(15-16), 4719-4742.