2023 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA) | 979-8-3503-0533-3/23/$31.00 ©2023 IEEE | DOI: 10.1109/ESMARTA59349.2023.10293480

Role of Big Data Analytics in The Internet of Things: A Comprehensive Survey

Mohammed A. Makarem
Department of Computer Engineering
King Fahd university of petroleum and minerals
Dhahran, Saudi Arabia
Makarem.a.Mohammed@gmail.com

Muneef A. Razaz
Department of Electriacl Engineering
King Fahd university of petroleum and minerals
Dhahran, Saudi Arabia
mneefbasha@gmail.com

***Abstract*— Due to the exponential proliferation of connected devices and the increase in data consumption rates, there are issues with collecting, processing, analyzing, and preserving massive amounts of Internet of Things (IoT) data. In recent years, there has been a significant increase in the number of internet-connected devices across various sectors. These devices include smart devices and sensors, which are utilized for data collection in diverse domains such as control systems, security measures, and healthcare applications which generate unstructured data. Consequently, the effective management of the IoT network necessitates the utilization of big data analytics. This investigation is meant to examine big data technologies and their various analytics platforms, as well as their significance within the IoT ecosystem through a comprehensive survey. This work aims to understand how big data analytics (BDA) and the IoT are related and the benefits of using platforms such as MAPR, Kaa, and Hadoop. Additionally, the study examines the requirements for analyzing big data in an IoT context. The research methods in this paper include comparing different approaches and platforms of BDA and IoT. The main finding is that security and privacy are major concerns when handling large amounts of data on the IoT.**

Keywords*— Internet of Things, Internet of Things in Big Data Analysis Platforms, massively parallel processing, High Availability Distributed Object-Oriented Platform, Big Data, and Artificial Intelligence.

I. Introduction

Big data refers to the vast amount of structured and unstructured data generated by various sources, including sensors, devices, and social media platforms. Big data analytics can provide valuable insights that can support decision-making and improve the efficiency of multiple processes. On the other hand, IoT refers to the interconnectedness of everyday devices, such as smart appliances and wearable devices, which can collect and transmit data in real time [1].

Cisco research claims that the IoT came into existence as a consequence of technical developments, the quick growth of wireless and electronic communications, and micro-electromechanical systems (MEMS) technologies [2]. The proliferation of Internet connectivity has resulted in a greater number of devices being interconnected than the total global population. In this research paper, Fig. 1 presents a comprehensive depiction of the diverse range of devices encompassed within the IoT. These devices encompass computers, cellphones, tablets, sensors equipped with Wi-Fi capabilities, and home appliances. The IoT facilitates the connection of these devices to the Internet, thereby enabling a multitude of functionalities and services. However, this interconnectedness also gives rise to complex challenges pertaining to data analytics. The present in Fig. 1 illustrates the existence of hybrid data sources in the context of IoT, which exhibit distinct characteristics and operate at exceptionally high data rates. Consequently, the field of IoT is regularly confronted with the challenges posed by massive amounts of data.

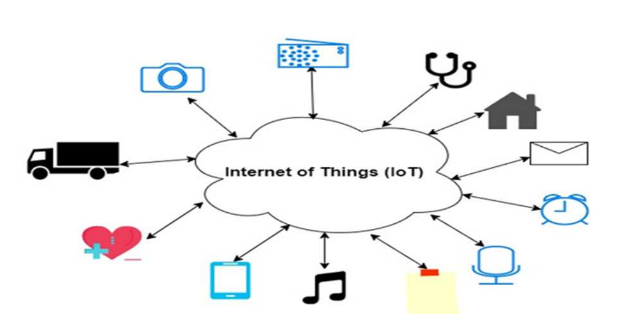

Fig. 1. The sources of Big Data in IoT.

According to research conducted by Siemens, it is projected that by the year 2020, approximately 26 billion physical objects will be interconnected on the internet. Also based on the findings of a report published by Statistica, it is projected that by the year 2025, the global count of interconnected devices will reach an estimated figure of 75.44 billion (refer to Fig. 2) [3].

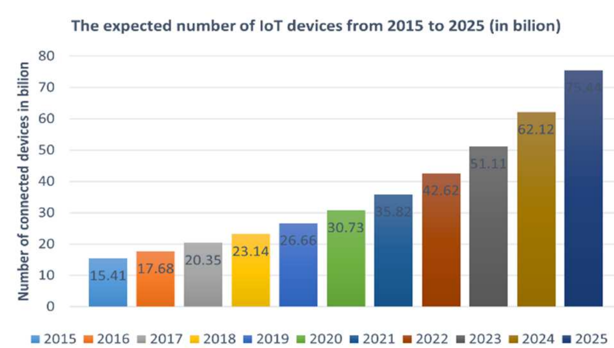

Fig. 2. Number of online devices from 2015 to 2025 in billion.

The majority of data-collecting devices used in the IoT environment include digital sensors that need the Message Queuing Telemetry Transport (MQTT) protocol and distributed

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

data services (DDS) [4]. Fig. 3 delineates the methodology of gathering data via IoT devices, monitoring the acquired data, and conducting data analytics.

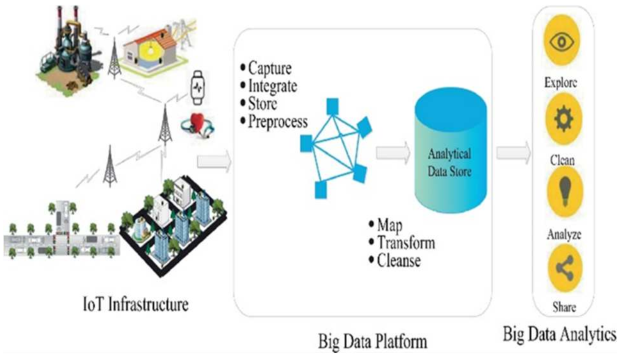

Fig. 3. The flow of big data analytics in the IoT environment [2].

Although the IoT has made it possible to save expenses and boost productivity, it cannot be fully used by just gathering enormous amounts of data. Institutions need to build a platform to handle and process the massive amounts of data that digital sensors generate. Information institutions may also use data analysis tools to convert enormous volumes of sensor data into usable information [5]. The complete forms of the abbreviations used in this paper can be found in TABLE 1.

TABLE 1 LIST OF ABBREVIATIONS.

Abbreviations	Definitions
MAPR	Map Reduce
Hadoop	High Availability Distributed Object-Oriented Platform
IoT	Internet Of Things
BDA	Big Data Analytics
MPP	Massively Parallel Processing
AI	Artificial Intelligence
MEMS	Micro-Electromechanical Systems
DDS	Distributed Data Services
API	Application Programming Interfaces
HER	Electronic Health Records
IoUT	Internet Of Underwater Things
RFID	Radio-Frequency Identification
IOPS	Input/Output Operations Per Second
PaaS	Platform As a Service
IPOL	Intelligent Data Operation
SQL	Structured Query Language
HDFS	Hadoop Distributed File System
SaaS	Software As A Service
JWT	Jason Web Tokens
QOS	Quality Of Services
IDC	International Data Corporation
ETL	extract-transform-load
MQTT	Message Queuing Telemetry Transport

The following are the key contributions identified in this paper:

- This paper aims to explore the topic of big data processing and analytics platforms within the context of the IoT environment. The discussion will revolve around the various platforms available for handling and analyzing large volumes of data generated by IoT devices. By examining the current landscape of big data platforms in the IoT domain, we hope to gain a deeper understanding of the challenges and opportunities associated with processing and analyzing data in this context.
- We will also delve into the essential prerequisites for the processing and analysis of large-scale data in the context of the IoT.
- We undertake a critical review of the recent literature.

In this section, general information was presented about what the world of IoT is expecting to be in the future by means of the number of connected devices, the sources of big data in IoT, and the flow of BDA in IoT. In addition, the main contributions of this paper are listed. The reset of the research paper is organized as follows:

In Section II, the arguments for choosing the investigation in this paper are presented. The study problem is stated in III. The study methodology used in this research is given in IV. In Section V, we review the existing research on big data and its role in the IoT. In VI, we critically compared some of the previous surveys on the same topic. In VII, we investigated the impact of big data on IoT. In Section VIII, the BDA' benefits in IoT applications were presented. In Section IX, the possible opportunities, and challenges by using BDA in the scope of IoT were investigated. In Section X, we listed the tools, techniques and platforms are being used within BDA and IoT. In Section XI, the requirements for using BDA in IoT are stated. In Section XII, we present some recent statistics about big data, the predominant subject areas for the application of BDA, and the mentioned platforms in Section X. We concluded this paper in section XIII. Finally, we provided some recommendations in section XIV.

II. ARGUMENTS FOR CHOOSING THIS INVESTIGATION

In this section, we present the arguments for choosing the investigation in this paper. These arguments justify the relevance and importance of our study and explain why we believe the topic we have chosen is worth investigating. The goal of this research is to clarify the relationship between BDA and the IoT.

A. Without analysis, the data produced by IoT devices is worthless.

B. Data volume increases with the advent of IoT applications.

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

C. Most of the data collected is unstructured and requires different processing and storage than traditional databases.

D. IoT and big data are mutually beneficial, and in order to gain insights from IoT applications, a data analytics platform is required to complement the two.

E. The convergence of IoT and big data is driving a technological revolution in the field of information technology, as both IoT and big data are current trends in our era.

III. THE STUDY PROBLEM

The number of Internet-connected devices, including computers, smartphones, Wi-Fi gadgets, digital sensors, and household appliances, has increased because of the IoT. In addition, the explosion of data volume, with terabytes now becoming petabytes, creates a more complex landscape for data analysis and processing. As a result, new methods of managing the enormous amount of collected data have been developed, and big data and artificial intelligence have been combined with other technologies to improve data management.

IV. STUDY METHODOLOGY AND TOOLS

This section provides an overview of the study methodology and tools used in our investigation. The methods used to conduct our research are important to understanding the validity and reliability of our findings. Through this section, we aim to provide a clear picture of the methods and tools used so that the reader can evaluate the rigor of our investigation and understand the justification for our choices. Furthermore, the study follows the approaches described in this section.

1. We employed an analytical approach to study open-source big data platforms and the IoT and their relationship to each other.
2. We employed a comparative approach to examine the relationship between BDA and IoT.

To accomplish this, we developed a checklist that included a set of standards and specifications for evaluating open-source IoT platforms. The criteria used in the evaluation were: integration, security and safety, databases, protocols, reliability, and support for device management and data analysis. Additionally, we looked at the impact of the IoT on big data, the challenges that arise from the intersection of the two, and the specific needs for BDA in an IoT context.

As illustrated in Table 2, a search of the Scopus database from 2016 to 2022 using the terms "IoT and Big Data" yielded approximately 110,637 results. The distribution of these results with respect to the number of each document type is also presented.

TABLE 2. CONTRIBUTING LITERATURE TYPE

Document Type	Number of Documents
Article	57940
Conference Paper	36433
Review	6567
Book Chapter	6263
Book	2391
Editorial	438
Conference Review	276
Note	112
Short Survey	111
Retracted	50
Letter	24
Data Paper	12
Erratum	11
Undefined	9

V. LITERATURE REVIEW

In this section, we review the existing research on big data and its role in the IoT. Integrating big data and IoT has led to new opportunities and challenges in various domains, such as healthcare, manufacturing, transportation, and smart cities.

This literature review will focus on the recent developments in the field of big data and IoT and their potential implications on the industry, businesses, and society. This literature review aims to identify the current state of the research on this topic, identify gaps in the literature, and provide a comparison of the surveyed papers. In addition, we examine the key concepts, architectures, technologies, and challenges related to big data and IoT and explore how they can be integrated to create new opportunities for innovation. We have divided the state-of-the-art studies into three categories:

1) Investigating BDA in the environment of the IoT

To address the challenges of storing and processing a significant quantity of data regarding smart buildings, Bashir et al.[6] suggested a framework for BDA and the IoT. Big data management, sensors, and the IoT are the three elements that comprise the proposed framework. To control the amounts of oxygen, smoke, and illumination in various areas of the smart building, the assessments are carried out in real-time. Cloudera Hadoop is implementing the framework, and all the analyses were created utilizing (Pyspark). The suggested framework has been expanded to include additional IoT applications, such as smart cities and smart airplanes, and is particularly created for smart buildings.

Moreover, a comprehensive examination is conducted by [7] to investigate the difficulties encountered in the implementation of BDA during the development of intelligent Industrial Internet of Things (IIoT) systems within the context of Industry 4.0.

In addition, [8] investigated advantages and disadvantages of the different traffic management systems and presented an inexpensive traffic control system by deploying IoT sensors for traffic detection every 500 or 1000 meters so to determine traffic density, the acquired data is sent to analytic programs. Not only that—but Zhang et al.[1] proposed a brand-new paradigm for computing that supports distributed data to be processed and shared in a collaborative environment based on the IoT firework model, which transforms distributed data physically by distributing virtual views of users' data. Moreover, Rathore et al. [9] suggested a smart city management system based on the IoT and BDA. Data is collected by deploying various sensors, including weather and water sensors, smart home sensors, and smart parking sensors.

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

The MapReduce Hadoop technology is used to implement the suggested scheme. Some of the procedures involved in the implementation process include data (production, collection, classification, and processing). The Spark over Hadoop technology is used to handle massive volumes of data effectively. Smart systems are used as data sources for the city to create a smart city as a functioning system. The constructed smart system hasn't been used yet; hence its correctness has not been evaluated.

Hidell et al. [10] discussed the importance of using the IoT to provide services to improve citizens' water, including air quality, transportation, and energy. The authors emphasize and produce a (Green IOT) system in Sweden to assess the benefits of using open-source software and open data to build smart cities. The study recommends that some guidelines related to the purchase of infrastructure should be developed open source and application programming interfaces API. Sezer et al.[11] suggested an improved framework that incorporates big data, the IoT, and semantic web technologies. The conceptual model of the IoT system is proposed in five layers: data acquisition, transformation, loading extract-transform-load (ETL), data extraction, logical reasoning, and learning. The data acquisition layer, which gathers data from many sources, may be thought of as the framework's input layer. Sensor drivers are provided by the ETL layer to convert the data gathered from various kinds of sensors. The resource description framework (RDF) data obtained from the ETL layer may be used by a reasoning engine to draw conclusions with the help of semantic-rule reasoning. The learning layer creates machine-learning-based models by extracting various properties from the data. The output of the learning layer is provided with specified actions by the action layer.

The work in [12] studied the role of big data in health care and found that digital sensors in the body generate a huge amount of health-related data. Two challenges are analyzed in this context: integrating big data with electronic health records (EHR) and presenting this data to doctors in real time.

Arora et al . [13] uses BDA techniques to classify devices that support the network, and the performance of four machine learning algorithms is analyzed. Also, Berlian et al . [14] established a framework for monitoring and downloading massive volumes of data produced by the marine and undersea IoT (IOUT), and the Map Reduce program is used to process this data. They used MapReduce to analyze the data and discover that, in comparison to SQL, the execution time of their queries is drastically reduced using this method. While this framework has numerous benefits, more work must be done in the future to evaluate the scalability of the scalable trust management protocol with IoUT applications and to establish trust-based admission control for IoUT systems.

According to research by Mourtzis et al. [15], the manufacturing sector has a lot of room to benefit from the introduction of IoT, which has the potential to completely replace antiquated processes with cutting-edge ones. Additionally, this process of transformation results in a data production-phase in which industrial data are transformed into industrial big data, which are meaningless without analytical capability. Businesses that embrace data analytics may discover novel, data-driven ways of dealing with market competition. They also show a simple business with over one hundred equipment how the Internet of Things concept may be put into practice.

The present energy development in India is analyzed by Gauret et al. [16] to assess the advantages that may be obtained by cloud computing and analytics. According to experts, the utilization of analytics has the potential to enhance operational efficiency, reduce operating costs, and empower customers by providing them with a greater degree of influence over the decision-making process. Lee et al. [17] presented an IoT-based electronic system that enables data analysis and knowledge acquisition approaches to boost productivity in diverse sectors. The authors provide a novel methodology for extracting big data from industrial informatics based on digital sensors, websites, and unstructured data. Moreover, GeeLytics is an edge analytics platform created by Papageorgiou et al. [18] that processes data in real time at the periphery of a network and in the cloud. This infrastructure caters to the needs of low-latency, geo-distributed analytics on massive volumes of IoT data. As a result of considering the system characteristics of heterogeneous edge/cloud nodes and the present workload, GeeLytics is built to accommodate a wide variety of dynamic stream processing topologies.

Wang et al. [19] discuss the problems and possibilities that the IoT and big data provide for the maritime group, as it establishes a new framework for connecting the industrial IoT with BDA, and its application helps to boost output and productivity. While Yen's et al. [20] explore how service discovery and composition methods may be used to address real-world issues using IoT data. They look at how technologies like data analytics and AI can be utilized in the smart world to get to the truth of the matter and respond appropriately. They propose a crowdsourcing platform built on the principles of online gaming to harness the power of human intellect in the execution of certain control tasks. In the foreseeable future, we will have mechanisms for proactive monitoring and diagnostics using a mix of big data.

In addition, an exploratory study of location privacy in the age of IoT, big data, and analytics is conducted by Minch et al. [21]. They pinpoint potential problems with BDA by identifying, categorizing, and describing privacy risks. They imply the need for the future creation of a dependable framework for protecting personal data in a situationally aware setting. Furthermore, Genoud et al. [22] conducted a survey to provide insight on the current approaches to dealing with big data and the problems that arise due to cyber-physical systems. Cloud security and the heterogeneous integration of data from many sources are at the heart of their research. They stress the need for real-time stream data processing and the necessity creating advanced data discovery systems.

Moreover, Paul et al. [23] suggested a system based on the IoT that can efficiently run parallel analytics on large amounts of data. Their goal is to find an appropriate analytical algorithm that can handle the demands of processing and evaluating massive data sets. Because of the great efficiency of the parallel

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

analytic algorithms in an IoT setting, their qualitative analysis yields encouraging outcomes. It is imperative that future research solve the challenges that prevent this paradigm from being put into practice in the context of fog computing. Also, Gao et al. [24] suggest a technique for clustering databases that may be used for big data analysis in the IoT paradigm (IOTStatisticDB). IOT-StatisticDB users input statistical functions using statistical operators inside the DBMS kernel. Multiple servers are used to do the statistical analysis in a decentralized and parallel manner.

2) Studies on BDA

Ahmed et al. [25] conducted a study dealing with the definition of big data and measuring associated trends in the field of information library management. The study aimed to analyze the research discovered in the knowledge network of Thomson Reuters. The study reached several conclusions, the most important of which are: reshaping the definition of big data and she noted that risk analysis, industry, and BDA are the most frequent research directions associated with big data. As well as, Rejeb et al. [26] studied the applications of the IoT by conducting a citation analysis of the Scopus database between 2011 and 2020.

Besides, this study [27] indicated that libraries provide many of their services through electronic means, which makes it necessary for him to use the big data that is collected from these means. The study also aimed to identify ethical issues related to the use of big data in libraries. The main goal of the study in [28] was to find out if 11 public institutions in the Netherlands had the organizational features needed to use big data. The study results, the most important of which was the existence of data systems operating under effective information technology that facilitated the sharing of information with other organizations. Additionally, the work in [29] indicated the importance of analyzing large volumes of data, and these large volumes of data cause problems with indexing and relational restrictions that slow down the performance of data storage and retrieval.

3) Studies on the IoT

Kamalrudin et al.[30] analyzed security requirements for IoT applications in libraries. This investigation demonstrated the creation of a new library of IoT security standards for the creation of IoT applications. The library's tools showed how to elicit security needs for each of the IoT features of certain business application domains using an industrial scenario. Also, Gao et al .[31] proposed an algorithm method based on RFID and Wi-Fi wireless technology . Additionally, Agrill et al. [32] point out that integrating security protocols into the components of the IoT is a challenge due to its limited resources as well as the potential for hardware damage from a few Internet users. Along with Rachit et al.[33] they pointed out that integrating security protocols into the components of the IoT is a challenge due to its limited resources as well as the potential for hardware damage from a few Internet users. Besides this study [34] analyzes the value of the IoT as the Internet of the Future and discusses the integration of IoT applications with artificial intelligence. Not only – but also Olson et al. [33] they worked on a study publication trend regarding the IoT, semantic web, and cloud computing based on the aggregate analysis of scientific articles, it was found that IoT applications ranked first in China. In addition, Fernandez et al.[35] introduced the concept of the "library of things," where he discussed the impact of IoT technology on libraries. Finally, Yao et al. [36] designed a smart system for energy-saving lighting in the library based on the IoT.

VI. COMPARISON WITH PREVIOUS STUDIES

In this section, we present a clear and comprehensive rationale for our research, highlighting its significance and the gap in current knowledge that our study aims to investigate by comparing it with some previous surveys on the same topic as summarized in TABLE 3.

In the previous years, academics have paid close attention to big data because of its enormous potential, as demonstrated by a database. Scopus Data in the past decade, 16038 articles related to big data have been published. In contrast, twelve years ago, there were only 26 articles published related to big data. Although there was significant research in both big data and IoT separately, research on the role of BDA in IoT still needs more investigation. In this section, we present a brief comparison between this survey and some previous related surveys within the same topic.

In [37], authors emphasize the significance of BDA and IoT in the context of disaster management. They delve into an examination of recent studies that have been conducted in this area. Additionally, the research in [38] was conducted to identify the role of BDA and IoT in supply chain management.

On the other hand, the authors in [39] have undertaken a comprehensive survey on the utilization of Big Data technologies within various domains of the Internet of Things (IoT). The primary objective of this survey is to promote and enhance the dissemination of knowledge pertaining to the application of big data technologies across different IoT domains.

Moreover, the research in [26] investigated big data and IoT applications by analyzing reference citations in the Scopus database from 2006 to 2015.

To our best of knowledge, the previous studies on BDA and the IoT are limited to specific fields, and the generalizations obtained from big data research and its relationship to the IoT are incomplete. This paper reveals the relationship between BDA and the IoT by providing critical facts and statistics about the different BDA platforms usage in the domain of IoT.

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

TABLE 3 A BRIEF COMPRASON BETWEEN OUR SURVEY AND PREVIOUS SURVEYS ON THE SAME TOPIC.

Reference	Year of publication	The focus of the paper	IoT platforms	BDA platforms	Provide statistics about BDA and IoT
[37]	2019	Highlights the importance of BDA and IoT for disaster management.	Yes	Yes	Not
[38]	2020	BDA and IoT in supply chain management.	Not	Not	Not
[39]	2018	Big data technologies in different IoT domains to facilitate and stimulate knowledge sharing across the IoT domains.	Not	Not	Not
The present survey	-	In this paper, we focus on how BDA and the IoT are related and the benefits of using platforms such as MAPR, Kaa, and Hadoop. Additionally, the study examines the requirements for analyzing big data in an IoT context.	Yes	Yes	Yes

VII. THE INFLUENCE OF BIG DATA ON THE IOT

IoT technology has made it possible to connect anything to the network. IoT links a network of people, things, apps, and data, allowing for remote control, integration, and management of services. As a result, this network is thriving since it requires a system for collecting and storing data created by IoT devices. As a result, new methodologies or engineering patterns in the domains of data gathering, storage, processing, and retrieval are required [40].

- Databases created and installed to operate with the IoT have requirements and characteristics, and the proliferation of NoSQL unstructured databases indicates that managing the IoT necessitates new database management approaches. In other words, fundamental need for big data storage is to be able to manage the enormous volumes of data and the input/output operations per second (IOPS) required to offer the data-to-data analysis tools, and the data center must be able to handle variable formats [41].
- By providing cloud computing platforms based on the IoT, possibilities to profit from its accomplishments and services are made more accessible [42].
- IoT has introduced new security concerns that conventional security measures cannot address [43]. Access control, data center security, redundant data security and filtering, secure accounts in a distributed environment, and real-time security enforcement are the security issues [44].
- Knowledgeable IoT data analysts and access to the wonderful world of big data will be required. BDA are urgently required for the IoT to help make decisions [45][46].

As a result, the IoT has an impact on people, processes, data, and things. First and foremost, there are people: more things can be observed and managed, increasing a person's capabilities; also, the operations of several devices will be used by the beneficiaries and interact with each other in real time; additionally, Data: the ability to collect large amounts of reliable data that leads to correct decisions and then use that data in decision-making includes, most recently, things like the ability to control things more accurately [40].

IoT affects our lives in major ways through a range of technologies such as SQL-structured databases in platform-as-data, Pig, Map Reduce, Sky Tree, Lambda, and Hadoop, which help in dealing with the huge amount of data generated by the

IoT. Some examples of the effect of IoT in our daily lives such as at work, a surveillance camera can estimate the time we spend there. How long you spend writing on the board in class may be determined by sensors. It could be a measure of employee productivity. At home a favorite movie is playing once the TV is on, smart devices can save a lot of energy and money by automatically turning off electrical appliances when leaving the house, and smart wrist devices linked to the elderly at home can contact the nearby hospital when patients need assistance [47] [48].

VIII. BDA' BENEFITS IN IOT APPLICATIONS

The majority of big data comes from IoT applications, and BDA is used in a variety of IoT applications, such as smart grids, healthcare systems, smart transportation systems, and smart inventory systems [49]. The advantages of BDA for IoT applications are shown in

Table 4. The classification of big data in IoT is shown in Table 5.

IX. OPPORTUNITIES AND CHALLENGES

In this section, the possible opportunities and challenges that will be created because of using BDA in the scope of IoT. BDA are made possible by the IoT ecosystem and the following opportunities:

A. Decision-Making

With the proliferation of IoT devices, smartphones, and social media, it is possible to gather useful data and forecast future trends. Big data may provide significant value and improve the usability and transparency of information in

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

information companies, allowing them to identify inconsistencies and improve performance [58].

B. *Efficiency Improvement*

Big data technologies like Hadoop can save a lot of money over traditional methods, and BDA can help reduce processing costs and make it easier to analyze data [59].

C. *Independence from the Locations of Data Storage*

Because of technologies like as big data, cloud computing, semantic web, and data warehousing that complement IoT. The emergence of IoT technologies has enabled the use of multiple data types, runtime data, device identification data, commercial data, retail data, and corporate data are all examples of such data [60].

D. *Value-added Applications*

Before the IoT and cloud computing, significant technologies like machine learning and artificial intelligence offered value-added applications and vast volumes of data. Since certain applications need enormous volumes of data, these technologies have not been used, but lately, data analysis platforms, business intelligence platforms, and analytical apps have emerged to aid information institutions in increasing their productivity [61] [62].

The IoT and big data provide difficulties. When it is addressed, it may have significant benefits in terms of big data quantities, difficulties in collecting data, inconsistent standards, new security concerns, data unreliability, a massive quantity of data to evaluate, and a quickly developing privacy environment [63]. Security, privacy, and data collection are among the most difficult problems we face. Together, BDA in the IoT environment faces challenges, the most important of which are the large number of IoT devices, the volume of data generated, the need to process and analyze the data, the need to provide reports, and the fact that the techniques are not standardized [64] [65].

TABLE 4. THE ADVANTAGES OF BDA FOR IOT APPLICATIONS [50].

IoT applications	Advantages of BDA
Intelligent transportation [51][52]	- Examining accident histories may help to reduce accidents. - Reduce traffic congestion. - Improve freight movements. - Assure traffic safety.
Smart healthcare [53]	- Prediction of epidemics, treatment, and patients. - Helping insurance companies develop better policies. - Recognizing anesthetic symptoms in the early stages of any major disorders.
Smart Grid [54][55]	- Assist in creating the best pricing strategy based on current energy use. - Projecting future supply requirements. - Assure a sufficient supply of power.
Intelligent management system [56]	- Develop strategic advertisements. - Recognizing client requirements. - Recognize possible dangers.

TABLE 5. CLASSIFICATION OF BIG DATA IN THE IOT ENVIRONMENT [57].

Source of big data	Components of the system	Possible technology for big data	Functional prerequisites	Analysis method
Management of smart cities	Data Gathering	Widespread Wireless Communication	The entry	Descriptive
Manufacturing	Data storage	Real time analytics	Predictive learning	Predictive
Information technology systems	Data Transfer	Machine Learning	Cloud Computing	Descriptive
Smart buildings	Data processing	Commodity sensors	Data extraction	-

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

X. Big Data and IoT Analytics Tools and Platforms

This section lists the tools, techniques and platforms are being used within BDA and IoT. There are several tools and techniques that are used to analyze big data, such as HPCC, Map Reduce, and Hadoop, and the most famous of these platforms is the Hadoop platform, which is open-source software, or a platform written in the Java language for storing and processing big data. Among the most famous users of Hadoop are Amazon, Apple, Facebook, Apple, LinkedIn, Twitter, and Yahoo. We will now see a comparison between all big data platforms and IoT platforms.

A. Big Data Platforms

- **Hadoop Apache** :An open source platform that provides storage, processing, and integration services between the various components of the IoT network and works using servers using MongoDB and HBase One of its benefits is that it can run any number of IoT applications on a single copy, and it also integrates with other platforms such as Ang Point and Mule, as well as using the MQTT and AMQP protocols. Grafana is also used to display data in graphical forms, and finally, H Base is used for non-relational databases [66].
- **1010data:** This tool offers analytics services and operates in a centralized way; nevertheless, the 1010Data platform is inefficient when it comes to data mining, loading, and transformation. It is a column database that primarily works with structured network data, such as IoT data [67].
- **Cloudera Data Center**: introduced the Enterprise Data Center, a framework for storing and analyzing IoT data built on the Hadoop platform that may serve as a focal point for handling enormous amounts of IoT data for businesses. One of this product's shortcomings is that it raises security and privacy issues [68].
- **SAP-Hana** is a big data and IoT analytics platform that accepts massive, unstructured data and supports the SAP protocol. The software examines both structured and unstructured data and uses Hive to access large data. In comparison to other tools, SAP's capabilities are not very strong [69].
- **HP-Haven platform** is a substantial new design for a universally applicable IoT data platform. In addition to search and discovery capabilities for unstructured data, the tool includes an Intelligent Data Operation Layer (IPOL), and Vertica provides a database management system for swiftly studying huge data. To aid in the discovery of huge data, HP also unveiled the Fex-Zone [70].
- **Pivotal platform** : Pivotal Analytics Database is divided into two types: Analytical Big Data (e.g. MPP) and ANSI SQL Structured Query Language. Pivotal Green Plum is an open-source database based on analytical big data [71] [72].
- **Infobright** : a device created primarily to address issues with data management and data analytics Up to 50 TB of data may be analyzed using Infobright suited for machine-generated data, such IoT data Large-scale data warehouses or the Hadoop platform are both used by Infobright [73].
- **Map Reduce:** BDA are supported by Map R, and more recently, Map R has introduced search, and Lucid wants Hadoop to have broadcast options. Map R offers more flexibility than Hadoop in order to improve its ability to forecast outcomes and speed up data processing [73].

B. IoT platforms

- **Sitwhere** was used first by Facebook, the MapReduce programming paradigm and the Distributed and File System (HDFS) are two of the most crucial parts of the Hadoop architecture. Yahoo is an open-source data storage and analysis system for enormous volumes of data. Its shortcomings include the inadequacy of tiny data sets [74].
- **The ThingSpeak Platform** allows the analysis and presentation of data within the MATLAB environment without the need to purchase a license. This platform provides the ability to collect and analyze sensor data in the cloud and develop IoT applications. It is compatible with Twitter's mobile and web apps. Among its advantages is data collection through special channels. It also uses the MQTT protocol [75].
- **DeviceHive** is an open, tool-rich platform distributed under the Apache License that is free to use, change, and modify and can be downloaded and used with private or public clouds. It is characterized by support for big data solutions such as Cassandra, Spark, and Apache, as well as support for libraries written in various programming languages and iOS and Android libraries [76].
- **Kaa** is a tool-rich platform that enables developers to manage databases for devices and related objects. Among its features are: managing an unlimited number of things; monitoring devices in real time; managing devices and adjusting their settings remotely; and finally providing smart A/B service testing services [77].
- **Zetta platform:** An open source platform for creating servers for IoT projects built using every node This platform can send data to data analytics platforms such as Spluk. This platform allows the integration of smartphone applications and hardware applications (things) with cloud applications in one place [78].
- **The Things Board platform:** A 100% open source platform for developing IoT applications, and you can add these applications to its cloud platforms as a service (SaaS) or as a product (Paas). This platform provides device management and data collection services and provides more than 30 tools that can be

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

built into user applications and also replace sensors and objects that are idle without turning off the system [79].

- **Thinger io:** An open-source platform for the IoT that gives connected devices a scalable cloud-based architecture. It can be controlled using the admin console interface. Among its advantages are cloud support, real-time control and monitoring panels, installing the server in your cloud, and using open libraries as a source for connecting devices [80].
- **WSO Platform**: An open source platform for IoT and smartphone applications that allows direct API use to connect your devices and integrates with boards in the market. provides edge computing and end-to-end processing capability through the WSO2 Siddh library. The platform supports HTTP and MQTT protocols. Self-service and management for iOS, Android, and Windows devices are supported. APIs support writing data visualization tools [81]. TABLE 6 also compares the specifications for open source IoT systems.

TABLE 6. SPECIFICATIONS FOR OPEN SOURCE IOT SYSTEMS.

Standards / platform	Manages hardware	integration	Security and confidentiality	protocols	analytics	Visualization support	Databases
Kaa	Yes	Integration of SDK components for each peripheral device that supports API applications	Encryption Link (SSL) RSA Key 204 bytes	MQTT COAP XMPP http TCP	Yes	Real-time Dante Kaa analytics and evolution with Apache Cassandra and Apache zeppelin support	MangoDB Hadoop SQL Oracle No
Thinger io	Yes	API support	SSL/data encryption	MQTT COAP HTTP higher reliability	WSO2 Data Analytics	Yes	Oracle my SQL or SQL Postgre
Device Live	Unknown	Apps support AQTT APIS	Trust using Jason web tokens (JWT)	Web Sockets MQIT	Apache Spark	Yes	SDA Hana Postgre SQL
Site where	Yes	Support for API applications	Link Encryption (SSL)	Support for MQTT AMQP STOMP Web Sockets protocol	Apache Sport Real-time Analytics	No	Mongo DB Li Base Influx DB
Thing board	Yes	API support	higher reliability	MQTT COAP HTTP	Apache Spark KAFKA Analytics	NO	Cassandra a
WSO2	Yes	API support	Data encryption higher reliability	HTTP MQTT WEB Socket COAP	Unknown	Yes	Cassandra a Mongo DB Influx DB SQL
Thing Speak	No	Apps support AQTT APIS	higher reliability	HTTP	View data within a Mata Lab environment	No	My SQL
Zetta	No	Application APIS support	higher reliability	HTTP	Splunk is used in non-real time	No	unknown

XI. REQUIREMENTS FOR BDA IN THE IOT ENVIRONMENT

In this section, the requirements for using BDA in IoT are stated. These requirements must be met within the IoT environment for efficiently utilize BDA.

A. Connection

One of the key goals of the IoT is to provide a consistent connection for big data and analytics in order to facilitate the integration of massive quantities of sensor data generated by the device. The IoT model gradually leads to universal communication between smart objects via digital sensors in a smart environment [82], and with the increasing presence of the Wi-Fi network and the 4G-LTE network for wireless Internet connection, it is already possible to see the development of information and communication networks. However, communication must be established between the various objects in smart cities [83], such as the IoT, cloud computing, and big data, as well as integrating information in a smart environment.

B. Storage

The fast proliferation of a vast number of things (the IoT) has resulted in the storing of massive volumes of data. Processing massive amounts of unstructured data from smart sensors and mobile devices, and social media, is one of the primary needs for storing big data in the IoT. They are built in various ways and use protocols and communication interfaces. It is diversified, and the majority of IoT services rely on MZM communication protocols, which are often used for cloud computing infrastructure [84].

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

C. Quality of services

The correct management of digital sensors, IoT, and smartphones is a need for Quality of Services (QOS) to evaluate enormous volumes of data in an accurate manner. Many studies have attempted to satisfy QOS standards; however unifying and integrating QoS with the IoT and big data needs further research and studies. The quality of service offered by IoT networks must be trustworthy, and it must contain an appropriate model for data from sources that create enormous amounts of data. Real-time IoT technology must be deployed to build a dependable network [85].

D. Real-time analytics

Making timely judgments now requires a vital effort called BDA. One of the most notable benefits of the Internet of Items is the ability to communicate information about "connected things" in real-time or near-real time. BDA should be used to assist information corporations make rapid choices and engage with people and technology in real time [86].

E. Standard

BDA has gotten a lot of interest from academics and other organizations. These businesses, however, confront considerable hurdles in storing and evaluating the enormous amounts of data generated by digital sensors in the IoT. This calls for the usage of BDA platforms, and the standard plays a significant role in this context by supplying enterprises with high-quality information for big data analysis [87].

XII. BDA AND THE IOT STATISTICS

In Section X, we briefly explained the different Big Data as well as IoT platforms by providing their specifications and technical usage. In this section, we present some recent statistics about big data, the predominant subject areas for the application of BDA, and the platforms in Section X.

Each day, more than 2.5 trillion bytes of data are created, accounting for 90% of the total data volume, according to IBM. Data from all across the globe [88] Facebook is used by one billion people every day, 31.25 million messages are received every minute; 2.77 million videos are viewed; and 300 hours of video are posted to YouTube. According to the International Data Corporation (IDC), the amount of data in the digital environment will double every two years. In 2025, the amount of data created by humans will exceed 181 zettabytes, or trillion gigabytes. This extraordinary increase in data leads to the formation of a data flood, also known as the data tsunami [89]. Fig. 4 shows how much the data has been exaggerated.

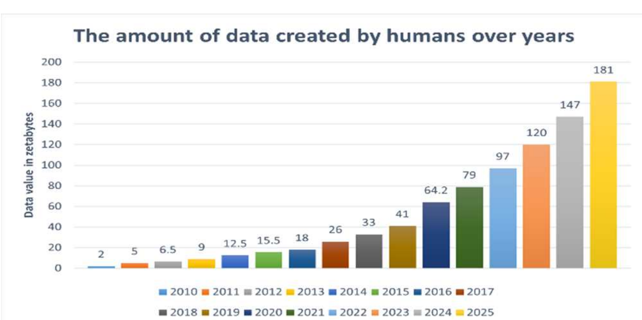

Fig. 4. How much the data has been exaggerated [2].

A. The IoT and Big Data Job Market

Big data and IoT are two of the most talked about technological subjects in recent years, and they are one of the key reasons why Gartner, an emerging technology analytics business, ranks them so highly. Furthermore, IDC predicts that more than half of the IoT activity would be centered in the domains of industry, smart transportation, and smart cities in 2010 recruitment for the IoT [90].The Central Journal of Data Science conducted a statistical study on the prevalence of IoT jobs today, and this is a list of the best IoT recruitment companies.

- *The best IoT recruitment firms:*
  1) *PTC*
  2) *Amazon*
  3) *Continental*
  4) *Savi Group*
  5) *Intel*
  6) *Ayla Network*
  7) *HP*
  8) *Log Meln. Inc Red Hat. Inc*
  9) *Honey Well*
  10) *Red Hat. Inc*
  11) *IBM*
  12) *Renesa*
  13) *Cisco Systems Inc*
  14) *Dell*
  15) *Inter Digital*

For the following jobs related to the IoT and big data, employers are seeking qualified candidates: IoT-related terms include "Big data, The Data World, Data Scientist, Digital Sensors, Sensor Data Engineer, and IoT Application Development".

a) Statistics about Big Data and the IoT

Within this subsection, an assortment of statistical data pertaining to the domains of big data and the IoT (IoT) is presented. Based on the analysis presented in Fig. 5, it is evident that the predominant subject areas for the application of BDA are primarily concentrated within the domains of computer science, accounting for 29.6% of the total usage, and engineering, accounting for 19% [28]. The utilization of big data extends beyond the realm of scientific disciplines, encompassing fields such as computer science and engineering, to encompass the social sciences, including accounting and business administration.

Fig. 6, which displays a statistical representation of the prevailing Big data platforms that are widely employed as analytical tools, shows that the vertical axes of the graph represent the rate of adoption of these systems in the market. In Fig. 7, the most popular IoT systems and their adoption rates in the market are represented. According to this statistic, the Hadoop platform is the most popular BDA platform, accounting for 90% of the market, followed by the Map R platform, accounting for 72.3%.

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

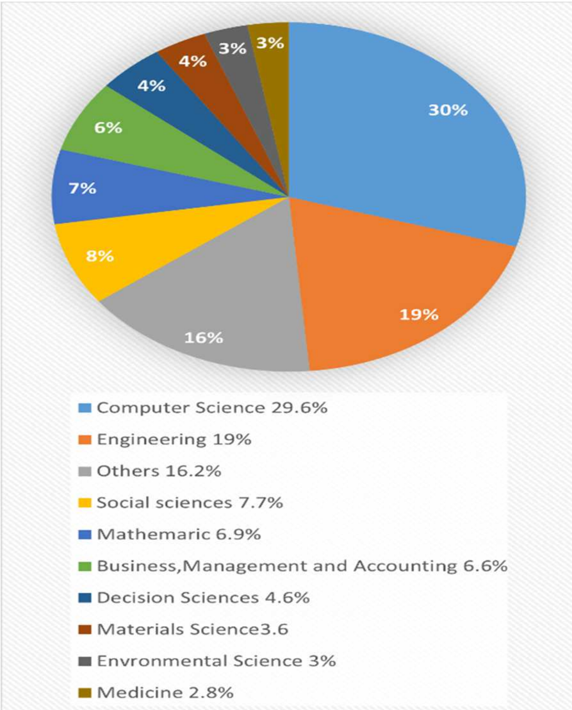

Fig. 5. Thematic distribution.

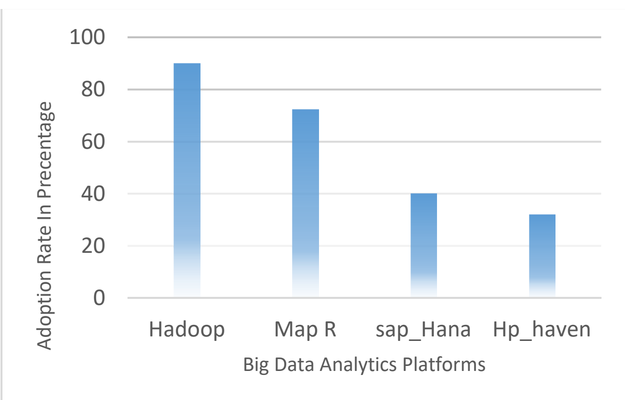

Fig. 6. Analytics systems for big data.

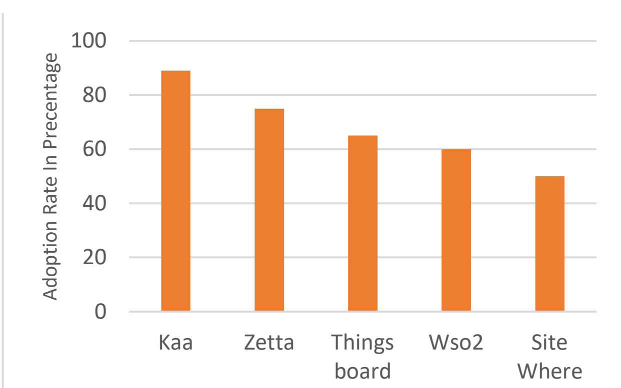

Fig. 7. IoT platforms.

Because of its incredible capacity to control devices, the Kaa platform is one of the most popular IoT systems, with an 89% adoption rate as shown in Fig. 7.

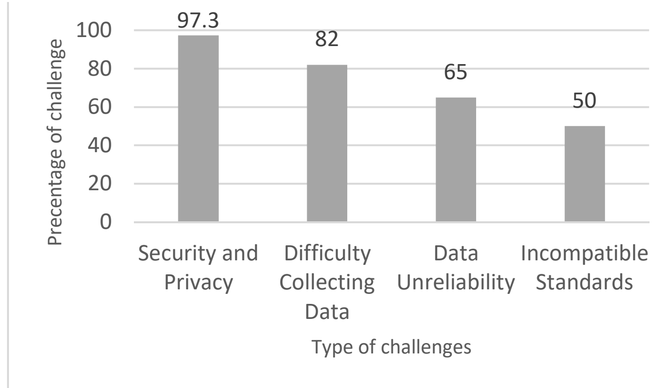

Fig. 8. The big data challenges with the IoT.

According to Fig. 8, one of the most significant issues with big data in the IoT is security and privacy, with a score of 97.3%, followed by data collection difficulty, with a rate of 82%.

XIII. CONCLUSION

This work is significant because it sheds light on the connection between IoT and BDA. Even though big data and the IoT are complementary, data analysis is used to accomplish this integration. The most prevalent security issues in IoT systems that leverage big data also include data centers, access management, and filtering redundant data. Instead, utilizing the IoT with big data may result in improvements in quality, safety, productivity, and decision-making via digital sensors. The IoT ecosystem also has the potential for big data to be used for decision-making, efficiency improvement, and value-added applications. In addition, with a percentage of 97.2%, security and privacy are the most significant problems confronting the IoT with big data, while data gathering comes in second with a proportion of 82%. Integration standards, security and confidentiality protocols, visualization capabilities, and databases are all included in the specifications for open source IoT systems. The two disciplines of computer science are among the subjects most often utilized for BDA, with a proportion of 29.6%, followed by engineering with a percentage of 19%. The Kaa platform is one of the most extensively used IoT systems because of its enormous capacity to control devices. Hadoop is also the most generally used BDA platform, with a rate of 90%, followed by Map R with a percentage of 72.3%. We concluded that some of the best IoT hiring companies are Amazon, IBM, PTC, Intel, Savi Group, and others. Lastly, real-time analytics, connection, storage, and service quality are some of the key factors for big data in the IoT.

XIV. RECOMMENDATIONS

Information firms must create platforms for handling and analyzing enormous amounts of digital sensor data, and data analysis technologies must be employed to transform a sizable quantity of data into actionable information. In addition to doing more studies on site privacy in the IoT and big data eras, it created principles for open-source infrastructure provision for

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

big data and the IoT through APIs. In addition, a better framework for integrating big data, the IoT, and semantic Web technologies is needed. Along with developing an electronic system based on IoT applications that supports data analysis and knowledge acquisition to increase productivity, other steps include offering qualified human cadres to work in big data management and proposing new technologies to collect and store data generated by the IoT. Finally, developing the necessary data analysis skills is necessary to deal with vast amounts of data.

XV. References

[1] Q. Zhang, X. Zhang, Q. Zhang, W. Shi, and H. Zhong, "Firework: Big data sharing and processing in collaborative edge environment," *Proc. - 4th IEEE Work. Hot Top. Web Syst. Technol. HotWeb 2016*, pp. 20–25, 2016, doi: 10.1109/HotWeb.2016.12.

[2] Indrajeet Bhattacharya, "Scientific Technology: Relevance and Role of 'Internet of Things (IoT)' in Museum," *J. Indian Museums*, vol. LXXII, no. 1, pp. 135–139, 2019, doi: 10.13140/RG.2.2.13764.37762.

[3] Statista Research Department, "Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025," *Stat. Res. Dep.*, vol. 44, no. 0, pp. 2024–2025, 2016.

[4] P. Mudholkar and M. Mudholkar, "Internet of Things (IoT) and Big Data: A Review," *Int. J. Manag. Technol. Eng.*, vol. 8, no. 12, pp. 5001–5007, 2018.

[5] F. J. Riggins and S. F. Wamba, "Research directions on the adoption, usage, and impact of the internet of things through the use of big data analytics," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2015-March, pp. 1531–1540, 2015, doi: 10.1109/HICSS.2015.186.

[6] M. R. Bashir and A. Q. Gill, "Towards an IoT big data analytics framework: Smart buildings systems," *Proc. - 18th IEEE Int. Conf. High Perform. Comput. Commun. 14th IEEE Int. Conf. Smart City 2nd IEEE Int. Conf. Data Sci. Syst. HPCC/SmartCity/DSS 2016*, pp. 1325–1332, 2017, doi: 10.1109/HPCC-SmartCity-DSS.2016.0188.

[7] Q. Qi, Z. Xu, and P. Rani, "Big data analytics challenges to implementing the intelligent Industrial Internet of Things (IIoT) systems in sustainable manufacturing operations," *Technol. Forecast. Soc. Change*, vol. 190, 2023, doi: 10.1016/j.techfore.2023.122401.

[8] P. Rizwan, K. Suresh, and M. Rajasekhara Babu, "Real-time smart traffic management system for smart cities by using Internet of Things and big data," *Proc. IEEE Int. Conf. Emerg. Technol. Trends Comput. Commun. Electr. Eng. ICETT 2016*, 2017, doi: 10.1109/ICETT.2016.7873660.

[9] M. M. Rathore, A. Ahmad, A. Paul, and S. Rho, "Urban planning and building smart cities based on the Internet of Things using Big Data analytics," *Comput. Networks*, vol. 101, pp. 63–80, 2016, doi: 10.1016/j.comnet.2015.12.023.

[10] B. Ahlgren, M. Hidell, and E. C. H. Ngai, "Internet of Things for Smart Cities: Interoperability and Open Data," *IEEE Internet Comput.*, vol. 20, no. 6, pp. 52–56, 2016, doi: 10.1109/MIC.2016.124.

[11] O. Berat Sezer, E. Dogdu, M. Ozbayoglu, and A. Onal, "An extended IoT framework with semantics, big data, and analytics," *Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016*, pp. 1849–1856, 2016, doi: 10.1109/BigData.2016.7840803.

[12] C. Vuppalapati, A. Ilapakurti, and S. Kedari, "The Role of Big Data in Creating Sense EHR, an Integrated Approach to Create Next Generation Mobile Sensor and Wearable Data Driven Electronic Health Record (EHR)," *Proc. - 2016 IEEE 2nd Int. Conf. Big Data Comput. Serv. Appl. BigDataService 2016*, pp. 293–296, 2016, doi: 10.1109/BigDataService.2016.18.

[13] D. Arora, K. F. Li, and A. Loffler, "Big data analytics for classification of network enabled devices," *Proc. - IEEE 30th Int. Conf. Adv. Inf. Netw. Appl. Work. WAINA 2016*, pp. 708–713, 2016, doi: 10.1109/WAINA.2016.131.

[14] M. H. Berlian *et al.*, "Design and implementation of smart environment monitoring and analytics in real-time system framework based on internet of underwater things and big data," *Proc. - 2016 Int. Electron. Symp. IES 2016*, pp. 403–408, 2017, doi: 10.1109/ELECSYM.2016.7861040.

[15] D. Mourtzis, E. Vlachou, and N. Milas, "Industrial Big Data as a Result of IoT Adoption in Manufacturing," *Procedia CIRP*, vol. 55, pp. 290–295, 2016, doi: 10.1016/j.procir.2016.07.038.

[16] R. Ramakrishnan and L. Gaur, "Smart electricity distribution in residential areas: Internet of Things (IoT) based advanced metering infrastructure and cloud analytics," *2016 Int. Conf. Internet Things Appl. IOTA 2016*, pp. 46–51, 2016, doi: 10.1109/IOTA.2016.7562693.

[17] C. K. M. Lee, C. L. Yeung, and M. N. Cheng, "Research on IoT based Cyber Physical System for Industrial big data Analytics," *IEEE Int. Conf. Ind. Eng. Eng. Manag.*, vol. 2016-Janua, pp. 1855–1859, 2016, doi: 10.1109/IEEM.2015.7385969.

[18] B. Cheng, A. Papageorgiou, F. Cirillo, and E. Kovacs, "GeeLytics: Geo-distributed edge analytics for large scale IoT systems based on dynamic topology," *IEEE World Forum Internet Things, WF-IoT 2015 - Proc.*, pp. 565–570, 2015, doi: 10.1109/WF-IoT.2015.7389116.

[19] H. Wang, O. L. Osen, G. Li, W. Li, H. N. Dai, and W. Zeng, "Big data and industrial Internet of Things for the maritime industry in Northwestern Norway," *IEEE Reg. 10 Annu. Int. Conf. Proceedings/TENCON*, vol. 2016-Janua, 2016, doi: 10.1109/TENCON.2015.7372918.

[20] I. L. Yen, G. Zhou, W. Zhu, F. Bastani, and S. Y. Hwang, "A Smart Physical World Based on Service Technologies, Big Data, and Game-Based Crowd

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

Sourcing," *Proc. - 2015 IEEE Int. Conf. Web Serv. ICWS 2015*, pp. 765–772, 2015, doi: 10.1109/ICWS.2015.111.
[21] R. P. Minch, "Location privacy in the era of the internet of things and big data analytics," *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, vol. 2015-March, pp. 1521–1530, 2015, doi: 10.1109/HICSS.2015.185.
[22] A. J. Jara, D. Genoud, and Y. Bocchi, "Big data for cyber physical systems an analysis of challenges, solutions and opportunities," *Proc. - 2014 8th Int. Conf. Innov. Mob. Internet Serv. Ubiquitous Comput. IMIS 2014*, pp. 376–380, 2014, doi: 10.1109/IMIS.2014.139.
[23] A. Mukherjee, H. S. Paul, S. Dey, and A. Banerjee, "ANGELS for distributed analytics in IoT," *2014 IEEE World Forum Internet Things, WF-IoT 2014*, pp. 565–570, 2014, doi: 10.1109/WF-IoT.2014.6803230.
[24] Z. Ding, X. Gao, J. Xu, and H. Wu, "IOT-StatisticDB: A general statistical database cluster mechanism for big data analysis in the internet of things," *Proc. - 2013 IEEE Int. Conf. Green Comput. Commun. IEEE Internet Things IEEE Cyber, Phys. Soc. Comput. GreenCom-iThings-CPSCom 2013*, pp. 535–543, 2013, doi: 10.1109/GreenCom-iThings-CPSCom.2013.104.
[25] W. Ahmed and K. Ameen, "Defining big data and measuring its associated trends in the field of information and library management," *Libr. Hi Tech News*, vol. 34, no. 9, pp. 21–24, 2017, doi: 10.1108/LHTN-05-2017-0035.
[26] A. Rejeb, K. Rejeb, S. Simske, H. Treiblmaier, and S. Zailani, "The big picture on the internet of things and the smart city: a review of what we know and what we need to know," *Internet of Things (Netherlands)*, vol. 19, 2022, doi: 10.1016/j.iot.2022.100565.
[27] D. Y. Qenawi, "The Role of Big Data Analytics in Internet of Things : Comparative Analytical Study," *Int. J. Libr. Inf. Sci.*, vol. 7, no. 2, pp. 74–111, 2020, doi: 10.21608/ijlis.2020.105260.
[28] B. Klievink, B. J. Romijn, S. Cunningham, and H. de Bruijn, "Big data in the public sector: Uncertainties and readiness," *Inf. Syst. Front.*, vol. 19, no. 2, pp. 267–283, 2017, doi: 10.1007/s10796-016-9686-2.
[29] B. Hussien Shahata Abdel Aziz, F. A. Omara, and D. HKhafagy, "Indexing For Improving Big Data Analysis," 2016.
[30] M. Kamalrudin, A. A. Ibrahim, and S. Sidek, "A security requirements library for the development of internet of things (IOT) applications," *Commun. Comput. Inf. Sci.*, vol. 809, pp. 87–96, 2018, doi: 10.1007/978-981-10-7796-8_7.
[31] Z. Gao, Y. Ma, K. Liu, X. Miao, and Y. Zhao, "An Indoor Multi-Tag Cooperative Localization Algorithm Based on NMDS for RFID," *IEEE Sens. J.*, vol. 17, no. 7, pp. 2120–2128, 2017, doi: 10.1109/JSEN.2017.2664338.
[32] A. S. Gillis, "What is IoT (Internet of Things) and How Does it Work? - Definition from TechTarget.com," *TechTarget*, 2022.
[33] Rachit, S. Bhatt, and P. R. Ragiri, "Security trends in Internet of Things: a survey," *SN Appl. Sci.*, vol. 3, no. 1, 2021, doi: 10.1007/s42452-021-04156-9.
[34] "A blockchain-and artificial intelligence-enabled smart IoT framework for sustainable city".
[35] P. Fernandez, "Through the looking glass: Thinking through the internet of things," *Libr. Hi Tech News*, vol. 32, no. 5, pp. 4–7, 2015, doi: 10.1108/LHTN-05-2015-0035.
[36] G. Y. Yao and L. Song, "Design of library lighting energy saving system based on internet of things," *Appl. Mech. Mater.*, vol. 496–500, pp. 1690–1693, 2014, doi: 10.4028/www.scientific.net/AMM.496-500.1690.
[37] S. A. Shah, D. Z. Seker, S. Hameed, and D. Draheim, "The rising role of big data analytics and IoT in disaster management: Recent advances, taxonomy and prospects," *IEEE Access*, vol. 7, pp. 54595–54614, 2019, doi: 10.1109/ACCESS.2019.2913340.
[38] A. Aryal, Y. Liao, P. Nattuthurai, and B. Li, "The emerging big data analytics and IoT in supply chain management: a systematic review," *Supply Chain Manag.*, vol. 25, no. 2, pp. 141–156, 2020, doi: 10.1108/SCM-03-2018-0149.
[39] M. Ge, H. Bangui, and B. Buhnova, "Big Data for Internet of Things: A Survey," *Futur. Gener. Comput. Syst.*, vol. 87, pp. 601–614, 2018, doi: 10.1016/j.future.2018.04.053.
[40] Z. Alansari, N. B. Anuar, A. Kamsin, S. Soomro, and M. R. Belgaum, "Computational intelligence tools and databases in bioinformatics," *4th IEEE Int. Conf. Eng. Technol. Appl. Sci. ICETAS 2017*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICETAS.2017.8277876.
[41] A. Kumar, S. A. Alghamdi, A. Mehbodniya, M. anul haq, J. L. Webber, and S. Navruzbek Shavkatovich, "Smart power consumption management and alert system using IoT on big data," *Sustain. Energy Technol. Assessments*, vol. 53, 2022, doi: 10.1016/j.seta.2022.102555.
[42] M. R. Belgaum, S. Soomro, Z. Alansari, S. Musa, M. Alam, and M. M. Su'Ud, "Challenges: Bridge between cloud and IoT," *4th IEEE Int. Conf. Eng. Technol. Appl. Sci. ICETAS 2017*, vol. 2018-Janua, pp. 1–5, 2018, doi: 10.1109/ICETAS.2017.8277844.
[43] S. Deep, X. Zheng, A. Jolfaei, D. Yu, P. Ostovari, and A. Kashif Bashir, "A survey of security and privacy issues in the Internet of Things from the layered context," *Trans. Emerg. Telecommun. Technol.*, vol. 33, no. 6, 2022, doi: 10.1002/ett.3935.
[44] M. Patyal, S. Sampalli, Q. Ye, and M. Rahman, "Multi-layered defense architecture against ransomware.," *Int. J. Bus. Cyber Secur.*, vol. 1, no. 2, pp. 52–64, 2017.
[45] B. N. Silva, M. Diyan, and K. Han, "Big Data Analytics," *SpringerBriefs Comput. Sci.*, pp. 13–30,

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

2019, doi: 10.1007/978-981-13-3459-7_2.

[46] P. G.-T. M. Review and U. 2014, "Big data characteristics," *Macrotheme.Com*, 2014.

[47] V. Saravanan, F. Hussain, and N. Kshirasagar, "Role of Big Data in Internet of Things Networks," pp. 273–299, 2019, doi: 10.4018/978-1-5225-7432-3.ch016.

[48] U. Ahmadova, M. Mustafayev, B. Kiani Kalejahi, S. Saeedvand, and A. M. Rahmani, "Big data applications on the Internet of Things: A systematic literature review," *Int. J. Commun. Syst.*, vol. 34, no. 18, 2021, doi: 10.1002/dac.5004.

[49] J. I. S. Jose, R. Zangroniz, J. J. De Dios, and J. M. Pastor, *Big Data and Internet of Things: A Roadmap for Smart Environments*, vol. 546. 2014.

[50] E. Ahmed *et al.*, "The role of big data analytics in Internet of Things," *Comput. Networks*, vol. 129, pp. 459–471, 2017, doi: 10.1016/j.comnet.2017.06.013.

[51] L. Zhu, F. R. Yu, Y. Wang, B. Ning, and T. Tang, "Big Data Analytics in Intelligent Transportation Systems: A Survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 1, pp. 383–398, 2019, doi: 10.1109/TITS.2018.2815678.

[52] S. Muthuramalingam, A. Bharathi, S. Rakesh kumar, N. Gayathri, R. Sathiyaraj, and B. Balamurugan, "Iot based intelligent transportation system (iot-its) for global perspective: a case study," *Intell. Syst. Ref. Libr.*, vol. 154, pp. 279–300, 2019, doi: 10.1007/978-3-030-04203-5_13.

[53] G. Manogaran, R. Varatharajan, D. Lopez, P. M. Kumar, R. Sundarasekar, and C. Thota, "A new architecture of Internet of Things and big data ecosystem for secured smart healthcare monitoring and alerting system," *Futur. Gener. Comput. Syst.*, vol. 82, pp. 375–387, 2018, doi: 10.1016/j.future.2017.10.045.

[54] V. Hashem, Ibrahim Abaker Targio Chang *et al.*, "The role of big data in smart city," *Int. J. Inf. Manage.*, pp. 784–758, 2016.

[55] Y. Zhang, T. Huang, and E. F. Bompard, "Big data analytics in smart grids: a review," *Energy Informatics*, vol. 1, no. 1, 2018, doi: 10.1186/s42162-018-0007-5.

[56] S. Paul, A. Chatterjee, and D. Guha, "Study of Smart Inventory Management System Based on the Internet of Things (Iot)," *IJRTBT Int. J. Recent Trends Bus. Tour.* |, vol. 3, no. 3, pp. 27–34, 2019.

[57] A. M. Vamsi, P. Deepalakshmi, P. Nagaraj, A. Awasthi, and A. Raj, "IOT Based Autonomous Inventory Management for Warehouses," *EAI/Springer Innov. Commun. Comput.*, pp. 371–376, 2020, doi: 10.1007/978-3-030-19562-5_37.

[58] J. M. Tien, "Internet of Things, Real-Time Decision Making, and Artificial Intelligence," *Ann. Data Sci.*, vol. 4, no. 2, pp. 149–178, 2017, doi: 10.1007/s40745-017-0112-5.

[59] S. P.-J. homepage: www. ijrpr. com ISSN, "A Survey on Big Data Analytics Data Science Technologies," *Ijrpr.Com*, vol. 2, no. 6, 2021.

[60] A. Bröring *et al.*, "Enabling IoT Ecosystems through Platform Interoperability," *IEEE Softw.*, vol. 34, no. 1, pp. 54–61, 2017, doi: 10.1109/MS.2017.2.

[61] J. Vidal-García, M. Vidal, and R. H. Barros, "Computational Business Intelligence, Big Data, and Their Role in Business Decisions in the Age of the Internet of Things," *Web Serv.*, pp. 1048–1067, 2018, doi: 10.4018/978-1-5225-7501-6.ch055.

[62] Y. Jeong, H. Joo, G. Hong, D. Shin, and S. Lee, "AVIoT: Web-based interactive authoring and visualization of indoor internet of things," *IEEE Trans. Consum. Electron.*, vol. 61, no. 3, pp. 295–301, 2015, doi: 10.1109/TCE.2015.7298088.

[63] A. Kobusińska, C. Leung, C. H. Hsu, S. Raghavendra, and V. Chang, "Emerging trends, issues and challenges in Internet of Things, Big Data and cloud computing," *Futur. Gener. Comput. Syst.*, vol. 87, pp. 416–419, 2018, doi: 10.1016/j.future.2018.05.021.

[64] H. N. Dai, H. Wang, G. Xu, J. Wan, and M. Imran, "Big data analytics for manufacturing internet of things: opportunities, challenges and enabling technologies," *Enterp. Inf. Syst.*, vol. 14, no. 9–10, pp. 1279–1303, 2020, doi: 10.1080/17517575.2019.1633689.

[65] S. Rizvi, A. Kurtz, J. Pfeffer, and M. Rizvi, "Securing the Internet of Things (IoT): A Security Taxonomy for IoT," *Proc. - 17th IEEE Int. Conf. Trust. Secur. Priv. Comput. Commun. 12th IEEE Int. Conf. Big Data Sci. Eng. Trust. 2018*, pp. 163–168, 2018, doi: 10.1109/TrustCom/BigDataSE.2018.00034.

[66] S. G. Manikandan and S. Ravi, "Big data analysis using apache hadoop," *2014 Int. Conf. IT Converg. Secur. ICITCS 2014*, 2014, doi: 10.1109/ICITCS.2014.7021746.

[67] V. Morabito, "Managing Change for Big Data Driven Innovation," *Big Data Anal.*, pp. 125–153, 2015, doi: 10.1007/978-3-319-10665-6_7.

[68] A. Bhardwaj *et al.*, "DataHub: Collaborative data science & dataset version management at scale," *CIDR 2015 - 7th Bienn. Conf. Innov. Data Syst. Res.*, 2015.

[69] F. Färber, S. Cha, J. Primsch, and C. Bornhövd, "SAP HANA database: data management for modern business applications," *ACM Sigmod …*, vol. 40, no. 4, pp. 45–51, 2012.

[70] S. Rouhani and S. Rotbei, "Big data platforms: in the lens of selection and evaluation approach," *J. Decis. Syst.*, 2021, doi: 10.1080/12460125.2020.1869432.

[71] Y. Arfat, S. Usman, R. Mehmood, and I. Katib, "Big data tools, technologies, and applications: A survey," *EAI/Springer Innov. Commun. Comput.*, pp. 453–490, 2020, doi: 10.1007/978-3-030-13705-2_19.

[72] Y. Zhuang *et al.*, "D-Ocean: an unstructured data management system for data ocean environment," *Front. Comput. Sci.*, vol. 10, no. 2, pp. 353–369, 2016, doi: 10.1007/s11704-015-5045-6.

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.

[73] F. Beer and U. Bühler, “In-database feature selection using rough set theory,” *Commun. Comput. Inf. Sci.*, vol. 611, pp. 393–407, 2016, doi: 10.1007/978-3-319-40581-0_32.
[74] J. Guth, U. Breitenbucher, M. Falkenthal, F. Leymann, and L. Reinfurt, “Comparison of IoT platform architectures: A field study based on a reference architecture,” *2016 Cloudification Internet Things, CIoT 2016*, 2017, doi: 10.1109/CIOT.2016.7872918.
[75] C. Acharya and S. Kuzhalvaimozhi, “Irrigation and Internet of Things Platform,” pp. 1643–1646, 2016.
[76] P. Ganguly, “Selecting the right IoT cloud platform,” *2016 Int. Conf. Internet Things Appl. IOTA 2016*, pp. 316–320, 2016, doi: 10.1109/IOTA.2016.7562744.
[77] L. Cao, “Comparison between SensibleThings and Kaa platform,” *Technology*, 2016.
[78] F. Ribeiro and J. Metrôlho, “An Internet of Things platform to facilitate the development of context aware applications. Overview , challenges and experiments,” *Int. J. Comput.*, vol. 1, pp. 104–111, 2016.
[79] A. Polianytsia, O. Starkova, and K. Herasymenko, “Survey of hardware IoT platforms,” *2016 3rd Int. Sci. Conf. Probl. Infocommunications Sci. Technol. PIC S T 2016 - Proc.*, pp. 152–153, 2017, doi: 10.1109/INFOCOMMST.2016.7905364.
[80] “EBSCOhost | 114795507 | Internet of Things in Learning Systems - A Perspective of Platforms.”
[81] K. Astropekakis, E. Drakakis, K. Grammatikakis, and C. Goumopoulos, “A Survey of IoT Software Platforms,” *Lect. Notes Networks Syst.*, vol. 289, pp. 299–326, 2022, doi: 10.1007/978-3-030-87049-2_10.
[82] Al Nuaimi Eiman, Al Neyadi Hind, Mohamed Nader, and Al-Jaroodi Jameela, “Applications of big data to smart cities,” *J. Internet Serv. Appl.*, vol. 6, no. 25, 2015.
[83] T. Saber, S. Miadreza, S. Pierluigi, L. Vincenzo, T. Aurelio, and C. João, “A Review of Smart Cities Based on the Internet of Things Concept,” *Energies*, vol. 10, no. 4, 2017.
[84] G. Suciu *et al.*, “Big Data, Internet of Things and Cloud Convergence – An Architecture for Secure E-Health Applications,” *J. Med. Syst.*, vol. 39, no. 11, 2015, doi: 10.1007/s10916-015-0327-y.
[85] J. Jin, J. Gubbi, T. Luo, and M. Palaniswami, “Network architecture and QoS issues in the internet of things for a smart city,” *2012 Int. Symp. Commun. Inf. Technol. Isc. 2012*, pp. 956–961, 2012, doi: 10.1109/ISCIT.2012.6381043.
[86] M. Objectives, “Real-Time IoT Stream Processing and Large-scale Data Analytics for Smart City Applications,” *Eu Fp7*, vol. 18, no. 609035, pp. 2–3, 2013.
[87] C. A. Ardagna, P. Ceravolo, and E. Damiani, “Big data analytics as-a-service: Issues and challenges,” *Proc. - 2016 IEEE Int. Conf. Big Data, Big Data 2016*, pp. 3638–3644, 2016, doi: 10.1109/BigData.2016.7841029.
[88] D. Gaurav, J. K. P. Singh Yadav, R. K. Kaliyar, and A. Goyal, “An Outline on Big Data and Big Data Analytics,” *Proc. - IEEE 2018 Int. Conf. Adv. Comput. Commun. Control Networking, ICACCCN 2018*, pp. 74–79, 2018, doi: 10.1109/ICACCCN.2018.8748683.
[89] Statista, “Total data volume worldwide 2010-2025 | Statista,” *Statista Research Department*, 2022.
[90] S. Nizetic, Solic. Peter, D. Artaza, and Patrono Luigi, “Internet of Things (IoT) Opportunities, issues and challenges towards a smart and sustainable future,” *J. Clean. Prod.*, 2020.

Authorized licensed use limited to: The National Institute of Engineering. Downloaded on December 12,2024 at 18:49:04 UTC from IEEE Xplore. Restrictions apply.