Iraqi Journal for Computer Science and Mathematics

Journal Homepage: http://journal.esj.edu.iq/index.php/IJCM
e-ISSN: 2788-7421 p-ISSN: 2958-0544

Unlocking the Potential: Synergizing IoT, Cloud Computing, and Big Data for a Bright Future

Zeena N. Al-kateeb[1*], Dhuha Basheer Abdullah [2]

[1,2] Department of Computer Science
College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

Corresponding author: Zeena N. Al-kateeb

DOI: https://doi.org/10.52866/ijcsm.2024.05.03.001
Received January 2024; Accepted March 2024; Available online June 2024

ABSTRACT: The convergence and integration of the Internet of Things (IoT), Cloud Computing (CC), and Big Data (BD) offer huge potential for transformative progress that will support the massive industrial revolution that is so prevalent in today's digital landscape. This prompts an exploration of the synergies among these three domains and investigates how their integration can unlock new opportunities for a brighter future. This integration seamlessly connects billions of IoT devices, leverages the potential of CC, and efficiently manages vast datasets generated by these devices. The amalgamation of BD analytics with IoT and CC empowers organizations to extract invaluable insights, foster data-driven decision-making, and fuel innovation across diverse industries. Nevertheless, the incorporation of these systems also gives rise to notable difficulties, encompassing issues such as data protection, worries about privacy, the ability to scale, and the management of data. The paper delves into these challenges, explores strategies, and examines the hurdles and best practices to address them effectively through a comprehensive examination of the potential benefits, challenges, and mitigation strategies. In addition, this paper offers insights into how synergizing IoT, CC, and BD can pave the way for a brighter, more promising future in technology and society.

Keywords: Internet of Things, Cloud computing, Big data.

1. INTRODUCTION

In the last few years, we have seen incredible developments in technology that have the potential to completely transform many facets of our lives [1]. These exciting new tech trends include integration between more than just techniques, such as artificial intelligence, virtual and augmented reality, the Internet of Things (IoT), and advanced robotics [1]–[2]. Among these trends, the integration of the IoT, cloud computing, and big data has garnered significant attention, which is considered a powerful combination that can shape a brighter future [2]. This integration brings together the vast network of connected devices in the IoT, the immense computing power and scalability offered by cloud computing, and the potential insights derived from analyzing large volumes of data through big data analytics [1]–[2].

The IoT, cloud computing, and big data analytics each offer tremendous value on their own. However, the real power lies in combining them to create an integrated ecosystem for data collection, storage, processing, and analysis [3]–[6]. The IoT provides a network of connected devices and sensors that generate vast amounts of diverse, dynamic data [7]. This data lacks structure and has overwhelmed organizations in the past. Now, cloud computing gives them the infrastructure and resources needed to efficiently store and manage that data [4]. Finally, big data analytics enables organizations to uncover hidden patterns and extract meaningful insights from this previously overwhelming IoT data [5]–[6]. It is through the synergy of these technologies that data can seamlessly flow through a connected ecosystem. This creates a valuable cycle of data collection, storage, processing, and analysis, generating the actionable insights organizations need to optimize operations, fuel innovation, and gain a competitive edge [3]–[7]. It is essential to leverage the cloud for real-time processing of sensor data, as this enables the exploitation of its potential. Such data processing enables the development of proactive and intelligent applications and services. Combining IoT with big data offers several benefits, including a deep understanding of the context and situation, real-time actionable insights, performance optimization, and proactive and predictive knowledge [8]. Cloud technologies provide decentralized and scalable information processing, analytics, and data management capabilities. In this paper, we introduce a cloud-based platform designed for IoT and big data applications, along with the corresponding prerequisites. This platform comprises several integral components, including a multitude of sensors and devices, big data analytics, cloud-based data management, edge-intensive computing, capabilities, and virtualization support [8]–[10].

*Corresponding author: zeenaalkateeb@uomosul.edu.iq

Nonetheless, the fusion of these three domains brings forth considerable challenges. Concerns regarding data security and privacy, scalability issues, and the efficient management of data represent formidable hurdles that require attention to fully unlock the potential of this integration [2]. Organizations must navigate these obstacles adeptly and institute robust strategies that guarantee the confidentiality, integrity, and accessibility of data while harnessing its complete potential for transformative outcomes [8].

This paper explores the complex interplay between IoT, cloud computing, and big data analytics. It examines their potential synergies and the transformational capabilities of integrating these technologies. Additionally, it evaluates how these components interact and elucidate the potential benefits and challenges they present. Furthermore, this paper aims to provide a comprehensive look at how organizations can pave the way for fully harnessing the combined might of IoT, cloud computing, and big data. By showing best practices to overcome key obstacles, we enable stakeholders to unlock the tremendous potential of these technologies while navigating the risks. With diligent examination of the benefits and challenges, we can work toward a future leveraging the full capabilities of IoT, cloud computing, and big data analytics in a secure, scalable, and responsible manner.

2. RELATED WORK

The study conducted by Nazari Jahantigh et al. [11] offers a systematic examination of the integration of the IoT and cloud computing, providing individual background insights into IoT and cloud computing before delving into their convergence. The selected papers (a total of 38) are categorized into applications, platforms, and integrations with other technologies, highlighting benefits such as scalability and real-time analytics. However, it does acknowledge persisting challenges, including standardization, security, privacy, and efficient resource management. The review, while valuable, lacks a comprehensive critical analysis of the strengths and weaknesses of the approaches discussed in the sample papers. Furthermore, the methodology for paper selection remains somewhat unclear, and the scope appears to be confined to recent years, disregarding important prior contributions. The conclusions, although broad, fall short of offering specific recommendations. In summary, this review addresses IoT-cloud integration but could benefit from a more in-depth exploration and a broader scope that includes a more comprehensive evaluation of prior literature. Chegini et al. [12] presented a comprehensive survey focusing on automation techniques within IoT-Fog-Cloud ecosystems. The exponential growth of connected IoT devices and the ensuing challenges related to big data and heterogeneity are aptly emphasized. The authors proposed six distinct types of automation functions geared towards mitigating these challenges, ranging from data filtering to workflow management. These automation strategies aim to enhance Fog offloading, accommodate heterogeneous components, and efficiently process extensive data volumes. Although the survey categorizes relevant literature according to these automation types, it falls short of a thorough critical analysis of the strengths and weaknesses of the techniques referenced. The scope primarily encompasses proposal-level works, neglecting studies evaluating real-world implementations and performance trade-offs, which could enhance the analysis. Selvaraj and Sundaravaradhan's review [13] delved into research trends and challenges surrounding IoT-enabled healthcare systems. It provides a comprehensive overview of IoT architecture, cloud integration, big data, and security mechanisms, with a specific focus on healthcare applications, including patient monitoring and assisting elderly patients. The benefits, such as remote healthcare access and early disease detection, are outlined, but limitations like high energy consumption, data security vulnerabilities, and scalability issues are also acknowledged. Despite covering a breadth of technologies and applications, the review lacks a critical assessment of the quality and comparative efficacy of the techniques it references. The scope primarily encompasses proposal-level work without substantial insights into real-world deployments and performance trade-offs. Additionally, potential solutions to the challenges raised are not discussed in detail [13]. Amin et al. [14] presented a systematic review of IoT technology advancements, emphasizing their integration with big data, data science, and network science. The paper explores how emerging approaches like machine learning, blockchain, and federated learning can enhance the next generation of IoT, particularly in areas such as architecture, data provenance, and edge/cloud computing. It identifies gaps, notably the need for a thorough evaluation of deep learning and federated learning in IoT contexts, and suggests that the research community prioritize these advanced methods to address limitations identified in previous literature. This review also provided a structured analysis of cutting-edge IoT enhancements, highlighting open issues and opportunities that align with findings from previous surveys [14]. In their comprehensive review, Ali et al. [15] addressed the enabling technologies and standards for IoT systems. IoT, as noted in prior surveys, faces critical challenges due to device heterogeneity, diverse underlying technologies, and the lack of standardization. This review adopted a layered architecture approach, examining the state-of-the-art and open issues at each layer of the IoT stack while also emphasizing the role of middleware platforms in facilitating IoT application development and integration. However, the review falls short in terms of in-depth critical analysis of real-world implementations and the comparative evaluation of the technologies referenced. The potential solutions discussed remain at a high level, lacking detailed technical insights. While it offers a broad overview of the IoT technology stack and middleware platforms, a deeper assessment of their practical efficacy could further enhance the analysis [15].

While the surveys by Nazari Jahantigh et al. [11], Chegini et al. [12], and Selvaraj and Sundaravaradhan [13] offer valuable insights into IoT, cloud computing, and automation techniques, they lack comprehensive critical analysis and implementation perspectives. The reviews by Amin et al. [14] and Ali et al. [15] provide broad landscapes of IoT advancements and technologies but do not delve into feasibility assessments. In contrast, the current paper provides an in-depth critical examination of the synergies between IoT, cloud computing, and big data analytics. The comparative assessment of various integration approaches and their efficacy in addressing key challenges is a notable strength absent in previous surveys. This work further stands out by furnishing actionable recommendations through best practices to overcome hurdles related to scalability, security, interoperability, and governance. The combination of conceptual frameworks, critical analysis grounded in technical rigor, and practical implementation guidelines makes this paper more comprehensive in unlocking the full potential of the IoT-cloud-big data nexus compared to prior literature reviews.

3. INTERNET OF THINGS (IoT)

The IoT has recently emerged as a prominent technological trend, finding diverse applications across sectors like energy, medicine, and agriculture [16]. Fundamentally, the IoT establishes connections and interactions between network-enabled devices [17]. In manufacturing, it is proving to be a powerful tool for control and automation by enabling remote devices to wirelessly communicate and share real-time data over the Internet [18]. The IoT refers to a network of interconnected devices that can collect and exchange real-time data within the network [16]–[18]. With capabilities to connect everything from industrial machines to wearable devices, the IoT allows remote monitoring, coordination, and automation, transforming traditional manufacturing processes. As this sophisticated technology continues to develop, the IoT is poised to revolutionize the manufacturing sector by optimizing systems, reducing costs, and enhancing quality control. The potential of IoT in manufacturing is just starting to be realized, and it represents an exciting frontier for innovation in the industry. This enables devices to gather and exchange data without the need for human intervention, utilizing network infrastructure and presenting numerous possibilities for integrating computer systems into the physical world [19]. IoT devices perform diverse functions, collecting and transmitting data to and from other devices within the network [20]. The IoT encompasses various smart devices such as TVs, mobiles, watches, glasses, and alarm monitoring systems, all connected through the Internet. It offers solutions for a wide range of applications, including health monitoring, smart cities, waste management, security, smart traffic, emergency services, retail, industry, and healthcare [21], as shown in Fig. 1.

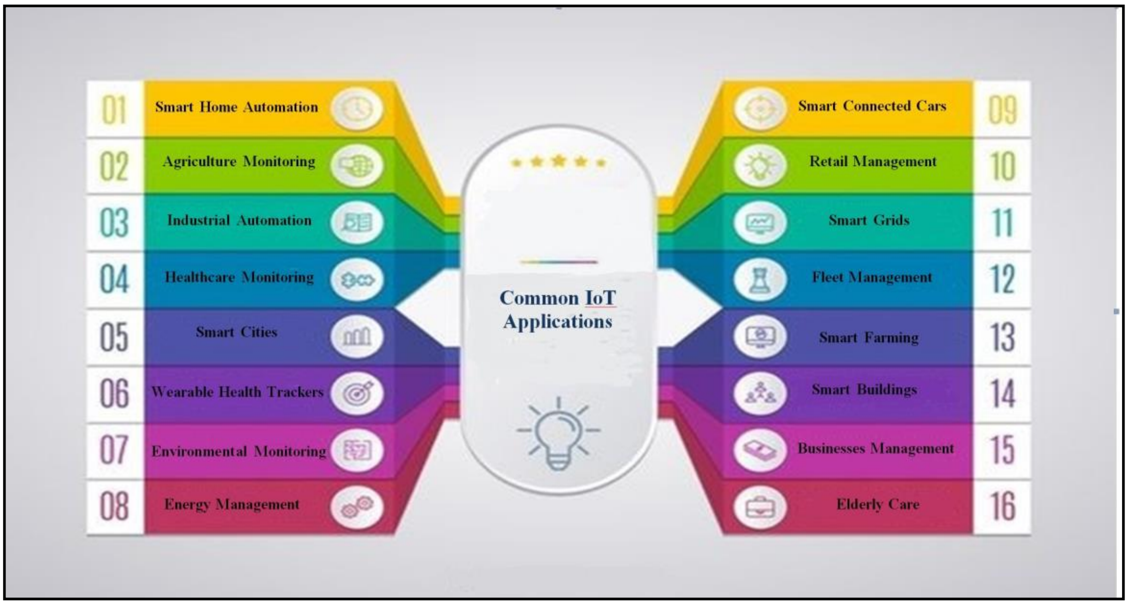

FIGURE 1. - Common IoT applications

4. CLOUD COMPUTING

Cloud computing refers to the idea of accessing resources conveniently and efficiently without the need to personally maintain the underlying hardware infrastructure [22], [23]. It encompasses a range of computing services, including essential databases, servers, crucial software, data analytics, and networking, all delivered over the Internet. This approach enables faster deployment, economic scalability, and the availability of flexible resources [24], [25]. Fig. 2 shows the importance of the facilities, features, and services provided by cloud computing.

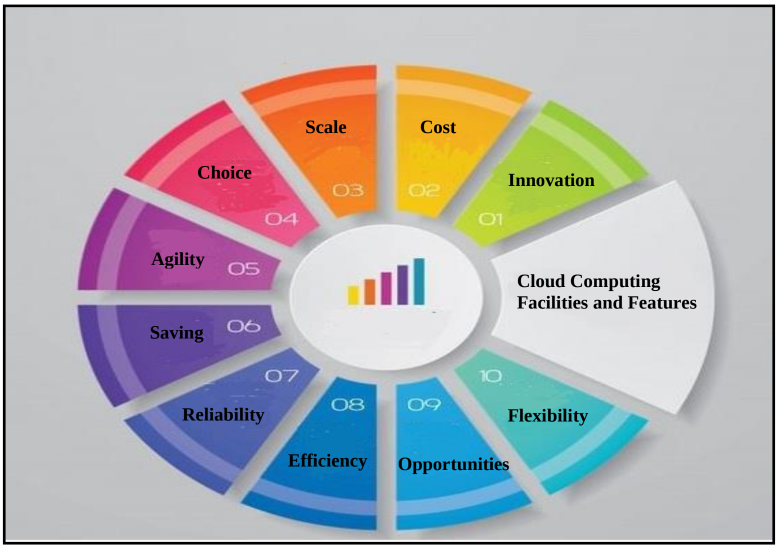

FIGURE 2. - **Cloud computing facilities and features**

5. BIG DATA

Big data refers to large volumes of structured, semi-structured, and unstructured data that cannot be easily processed or analyzed using traditional data processing methods. It is characterized by the three main V's: volume, velocity, and variety [26].

1. Volume: Big data involves vast amounts of data generated from various sources, including social media, sensors, transaction records, and more. The data can range from terabytes to petabytes or even exabytes in size [26], [27].
2. Velocity: Big data is generated at high speeds and requires real-time or near-real-time processing. The data is produced continuously, and timely analysis is necessary to extract meaningful insights and make informed decisions [26], [27].
3. Variety: Big data encompasses diverse data types, such as text, images, videos, audio files, log files, and structured data from databases. It includes both structured data (e.g., data in rows and columns) and unstructured data (e.g., social media posts and emails) that do not fit into traditional data models [26], [27].

In addition to these three V's of big data, there are expanded versions that include additional V's to further characterize the challenges and opportunities associated with big data. These expanded versions are the 5 V's, 7 V's, and 11 V's of big data. Let's explore some of them [28].

In the 5 V's of big data, the following two characteristics are added into three main common V's:

4. Veracity: This refers to the uncertainty, noise, and inconsistency that may exist within the data. Veracity focuses on the quality, reliability, and trustworthiness of the data, as data sources may be unreliable or contain errors.
5. Value: This signifies the ultimate goal of big data analysis, which is to extract actionable insights and create value for businesses, organizations, and society as a whole. The value derived from big data can be in the form of improved decision-making, innovation, cost savings, and competitive advantage.

The 7 V's build upon the 5 V's and two added additional dimensions:

6. Variability: This refers to the dynamic and changing nature of big data. Data can have fluctuations in volume, velocity, and variety, making it challenging to consistently handle and analyze [29].
7. Visualization: Visualization focuses on the importance of effectively presenting and interpreting big data. It involves using visual representations, such as charts, graphs, and dashboards, to make complex data more understandable and facilitate decision-making.

The 11 V's further expand on the characteristics of big data, including validity, volatility, viscosity, vicinity, vocabulary, volition, and value-to-volume ratio [29].

The expanded versions of the V's help provide a comprehensive understanding of the complexities involved in big data, emphasizing aspects beyond volume, velocity, and variety, such as veracity, value, and visualization. Fig. 3 illustrates the evolution of big data characteristics.

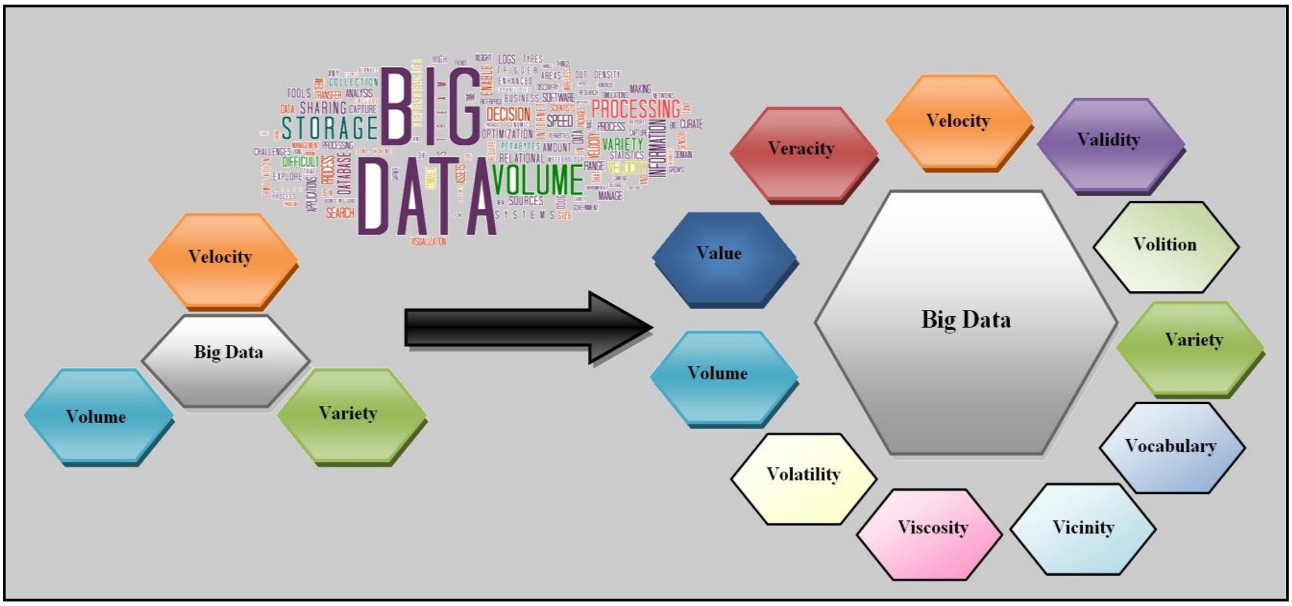

FIGURE 3. - Evolution of big data characteristics

6. COMPARATIVE ANALYSIS OF IOT, CLOUD COMPUTING, AND BIG DATA

While all three technologies, namely the IoT, cloud computing, and big data, are significant in their own right and have diverse applications, they exhibit fundamental distinctions [25]. These dissimilarities arise from variations in elements such as account management, communication protocols, accessibility, displacement characteristics, storage capacity, relationship with big data analytics, sensing capabilities, security and privacy considerations, availability, the role of the Internet, cost implications, applications, real-time processing capabilities, data size, data diversity, data speed, data validity, data integration, data monitoring, visual display of information, machine learning integration, fault tolerance mechanisms, and data privacy concerns. Table 1 provides a detailed comparison between the three techniques from various aspects.

Table 1. - Comparison between IoT, cloud computing, and big data

Criteria	IoT	Cloud Computing	Big Data
Scope	Limitless	Limited	Limitless
Components	Objects in the real world	Virtual resources	Large-scale datasets
Computation	Limited processing	Virtually limitless	Data-intensive processing
Communication	Device interaction	Data and application sharing	Data transfer and analysis
Reachability	Restricted	Ubiquitous	Ubiquitous
Displacement	Pervasive	Centralized	-
Storage	Limited	Long-lived storage	Large-scale data storage
Relation to big data	Data source	Data processing	Data utilization
Sensing Capabilities	Broad range of sensors	-	-
Security and privacy	Critical issue	Satisfies many requirements	Privacy concerns
Availability	Restricted	Distributed	-
Internet Role	Converging place	Delivery of service	-
Cost	Varies	Pay-as-you-go model	-
Applications	Smart homes, healthcare,	Business applications,	Predictive analytics,
Real-time Processing	Real-time data analytics	Real-time processing	Real-time analytics
Data Volume	Moderate	Variable	Massive
Data Variety	Sensor data, structured and unstructured data	Variable	Structured, unstructured, semi-structured data
Data Velocity	Moderate	Variable	High-velocity data streams
Data Veracity	Data from the physical world, potential for inaccuracies	Data reliability and integrity	Data accuracy and quality

Data Integration	Integration of diverse data sources	Data integration and consolidation	Data integration from multiple sources
Data Governance	Limited governance policies	Data governance and compliance	Data governance and regulatory compliance
Data Visualization	Customized for specific applications	Visualization tools and dashboards	Advanced visualization techniques
Machine Learning	Embedded machine learning algorithms	Machine learning capabilities	Machine learning for predictive analytics
Fault Tolerance	Susceptible to device failures	Redundancy and fault-tolerant infrastructure	Fault tolerance mechanisms
Data Privacy	Privacy concerns and data protection	Data privacy and security measures	Privacy implications and anonymization techniques

7. SYNERGISTIC INTEGRATION OF IOT, CLOUD COMPUTING, AND BIG DATA

Recognizing the distinct advantages and constraints of the IoT, cloud computing, and big data, there is a growing consensus on the value of amalgamating these technologies to harness their joint capabilities. Numerous integration scenarios have arisen, underscoring the advantages yielded by their collaborative utilization. These integration scenarios span a range of aspects, including:

Integration of IoT and Cloud Computing: The IoT generates massive volumes of real-time data from connected devices and sensors. By integrating IoT systems with cloud computing, organizations gain the scalable infrastructure needed to handle this data deluge. The cloud provides flexible, on-demand computing power for ingesting, processing, analyzing, and storing IoT data as it is produced. This real-time data flow enables advanced monitoring, automation, and predictive capabilities powered by cloud-based analytics and machine learning tools. The integration of IoT and cloud computing paves the way for organizations to harness the full disruptive potential of IoT-generated big data.

Integration of Cloud Computing and Big Data: The combination of cloud computing and big data presents an economical and scalable approach to store, process, and analyze extensive datasets. Cloud platforms furnish the essential infrastructure and distributed computing frameworks required for the efficient handling of big data workloads. This integration empowers organizations to unveil patterns, correlations, and trends within the immense data volumes, thereby facilitating informed decision-making and the deployment of predictive modeling and machine learning applications.

Integration of IoT, Cloud Computing, and Big Data: The intersection of IoT, cloud computing, and big data represents a pivotal transformational prospect for enterprises. By integrating all three technologies, organizations can collect, store, process, and analyze real-time data from IoT devices at scale. This integration allows for comprehensive data analytics, empowering organizations to extract valuable insights, optimize operations, and create innovative services. It also enables the deployment of advanced machine learning algorithms and predictive models, leading to intelligent automation and improved business outcomes.

7.1 IAAS, PAAS AND SAAS MODELS

A cloud-based IoT platform refers to a resource-sharing platform that offers dynamic and adaptable IoT services. It aligns with the NIST definition of cloud computing, which describes it as a model that provides convenient access to a shared pool of configurable computing resources [30], [31]. This model enables organizations to swiftly provision and release resources with minimal management effort or interaction with service providers. The cloud-based IoT platform operates through three service models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Its primary objective is to facilitate the utilization of IoT data by providing scalable resources and efficient service management. The platform ensures that the required resources are readily available for processing and analyzing massive amounts of IoT data. Additionally, it allows for resource scaling when needed, optimizing resource allocation and efficiency [30], [31].

The cloud-based IoT platform typically needs to choose between three service models: IaaS, PaaS, or SaaS. IaaS involves outsourcing computer infrastructure to support enterprise operations using resource virtualization. The success of cloud technology initially stemmed from its ability to host legacy applications within virtual machines (VMs) managed by external parties. This relieved application owners from the burden of managing physical infrastructure [32].

On the other hand, PaaS offers a platform for customers to develop, run, and manage applications, removing the complexities associated with infrastructure development and maintenance. PaaS frameworks compile applications from their source code and deploy them within lightweight VMs or containers. PaaS environments also provide tools for application scaling and task scheduling [31], [33].

In contrast, SaaS follows a software licensing and delivery model where software is centrally hosted and licensed on a subscription basis. It is accessed via a web browser and is often referred to as “on-demand software.” Cloud-based IoT platforms typically adopt the SaaS model, providing IoT-related services through a web interface on a pay-per-use basis. However, SaaS IoT platforms have limitations in terms of customizability and complex application development. While some offer extensibility mechanisms for user-provided callbacks, the resulting applications may lack homogeneity and become difficult to maintain [34].

With the generation of a huge amount of data from IoT, there is a need to organize the vast amounts of data generated today. Not only is there a significant increase in big data, but there is also a growing demand for data analytics platforms such as Hadoop. Consequently, this trend has opened up new possibilities in the field of cloud computing that is merged with IoT [35]. As a result, major service providers like Amazon Web Services (AWS), Google, and Microsoft have introduced their own cost-effective big data systems that can be easily scaled to accommodate businesses of all sizes. This development has given rise to a novel service model known as Analytics as a Service (AaaS). AaaS offers a swift and scalable approach to integrating various types of structured, semi-structured, and unstructured data. It enables real-time analysis, transformation, and visualization, providing businesses with efficient means to harness the power of data [33]–[35]. Fig. 4 shows the basic cloud computing models and models after integrating cloud computing with the IoT and big data.

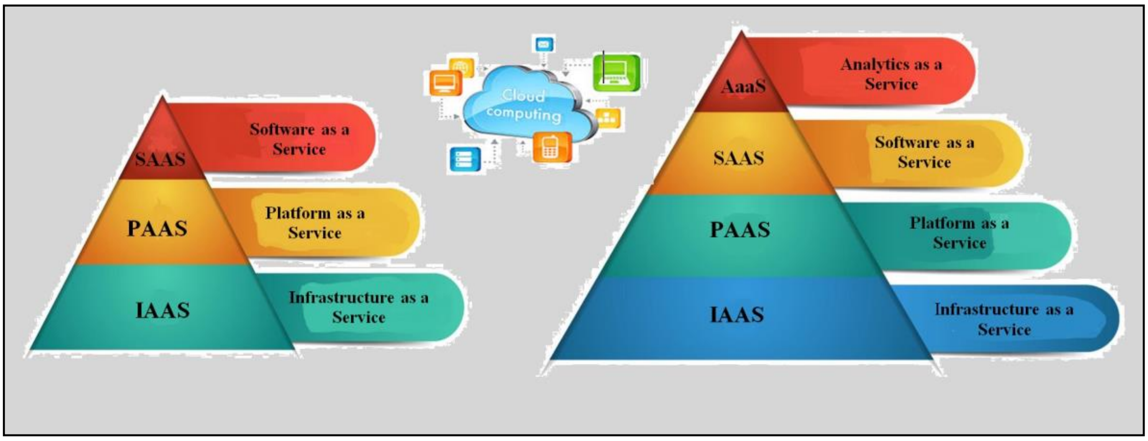

FIGURE 4. - **Cloud computing models and models before and after integrating with the Internet of Things and big data**

7.2 REQUIREMENTS OF THE IOT BIGDATA ANALYTICS PLATFORM

In constructing an IoT big data analytics platform, a host of crucial requirements come into play. Firstly, it must demonstrate the ability to dynamically manage IoT data, addressing the inherent challenges of connectivity among diverse and heterogeneous devices while tackling interoperability issues. A fundamental aspect is the platform’s capacity to glean valuable insights and knowledge from the vast reservoirs of IoT data at its disposal.

Furthermore, the platform should exhibit universality in accessibility and connectivity for objects, services, and users, with a particular focus on mobile contexts. In parallel, it should possess the agility required for seamless user management and orchestration, effectively handling the multitude of connected devices and the extensive data they generate. Finally, it should be geared towards user and service personalization, capable of delivering tailor-made services that align precisely with user preferences and real-world contexts.

The requirements for an IoT big data analytics platform can be summarized as follows:

- Intelligent and Dynamic Capabilities: The platform should incorporate intelligent and autonomic features to dynamically manage IoT functions, components, and applications. Additionally, it should be capable of making proactive decisions, dynamically deploying resources, and making intelligent choices based on an understanding of environmental context, user requirements, and application needs. This involves workload offloading from clients/hosts to the cloud and the dynamic migration of resources and services.
- Distributed Architecture: The platform should feature a distributed structure that encompasses various essential components. These include distributed information processing and computing capabilities, distributed storage, distributed intelligence, and distributed data management functionalities. These critical functions should be distributed across a network comprising smart devices, gateway/server systems, and multiple cloud environments. An important aspect of this architecture involves moving processing capabilities closer to end-users to minimize bandwidth usage. This distributed setup optimizes the platform’s performance and enhances its efficiency in managing and processing data across various nodes and environments.

- Scalability: The platform must possess the ability to scale effortlessly, accommodating varying numbers of devices, services, and users. This entails dynamically adjusting data management, storage, and processing services to meet evolving demands.
- Real-Time Processing: A core requirement is the platform's capability to process data in real-time, ensuring swift analysis and responses, particularly in urgent situations. Prioritizing time-sensitive tasks over non-urgent ones is imperative for effective real-time data analysis.
- Programmability: The platform should offer robust support for programmable features, allowing for the customization of IoT business and service logic, data warehousing schemes, and service models. Flexibility in customization and adaptation to specific requirements is a must.
- Interoperability: Seamless interoperability is an important fundamental requirement among diverse IoT services and infrastructures. This necessitates the use of APIs adhering to established standards and the publication and maintenance of components as open-source software. Where conceded, the ultimate objective is the establishment of a common data model capable of harnessing both structured and unstructured data. As well as the transition from raw data to linked data and the adoption of clear, unambiguous descriptions of pertinent information; these are essential steps in facilitating the creation of versatile, cross-domain smart applications.
- Security: The platform should prioritize security and privacy by design. This includes features such as data integrity, localization, confidentiality, and service level agreements (SLAs). Holistic approaches are necessary to address privacy and security concerns across value chains.

8. CONTRASTING IOT ARCHITECTURES: CLOUD AND BIG DATA VS. TRADITIONAL IOT

The IoT based on cloud and big data differs from traditional or non-cloud and big data-based IoT in several notable aspects:

- Scalability: Cloud-based IoT with big data allows for virtually unlimited scalability compared to traditional IoT systems. Cloud infrastructure can dynamically allocate resources based on demand, enabling the handling of large-scale deployments and massive amounts of data generated by IoT devices.
- Storage Capacity: Cloud-based IoT with big data leverages scalable and distributed storage systems, such as data lakes or distributed file systems, to accommodate vast amounts of data. Traditional IoT systems often have limited storage capacity and may require manual data management or archiving.
- Processing Power: Cloud-based IoT with big data can harness the computational power of cloud infrastructure for data processing, analysis, and machine learning tasks. Traditional IoT systems typically have limited processing capabilities and may rely on less powerful edge devices for basic data processing.
- Real-time Analytics: Cloud-based IoT with big data enables real-time or near-real-time analytics on the incoming data streams. By leveraging big data processing technologies like Apache Spark or stream processing frameworks, actionable insights can be derived promptly. Traditional IoT systems may lack real-time analytics capabilities, relying on batch processing or delayed analysis.
- Flexibility and Adaptability: Cloud-based IoT with big data provides flexibility in terms of data formats, protocols, and device integration. Cloud platforms often offer a wide range of APIs and tools for seamless integration with various IoT devices and protocols. Traditional IoT systems may be limited in terms of compatibility and interoperability with diverse devices and data sources.
- Cost Efficiency: Cloud-based IoT with big data offers cost advantages through pay-as-you-go pricing models. Organizations can scale resources as needed and avoid the upfront costs of establishing and maintaining an on-premises infrastructure. Traditional IoT systems require significant upfront investments in hardware, maintenance, and infrastructure setup.
- Data Security and Privacy: Cloud-based IoT with big data emphasizes robust security measures, encryption, access control, and data anonymization. Cloud service providers typically adhere to industry best practices and compliance regulations. Traditional IoT systems may have vulnerabilities due to limited security measures or may require additional effort to ensure data privacy.
- Integration with Analytics Ecosystem: Cloud-based IoT with big data easily integrates with existing analytics ecosystems and tools. Organizations can leverage the rich analytics and machine learning capabilities provided by cloud platforms or integrate with third-party analytics services. Traditional IoT systems may lack seamless integration with advanced analytics tools, requiring additional effort for analysis and insight extraction.
- Maintenance and Updates: Cloud-based IoT with big data reduces the burden of system maintenance and updates. Cloud service providers handle the underlying infrastructure, ensuring availability, security patches, and software updates. Traditional IoT systems require dedicated resources for maintenance, updates, and ensuring system availability.

IoT based on cloud and big data offers enhanced scalability, storage capacity, processing power, real-time analytics, flexibility, cost efficiency, security, integration, and simplified maintenance compared to traditional/non-cloud and non-big data-based IoT systems. Table 2 provides a general overview of the differences and may not capture all possible variations in IoT implementations.

Table 2. - Differences between the Internet of Things based on cloud and big data and traditional IoT

Pect	IoT based on Cloud and Big Data	Traditional IoT
Scalability	Virtually unlimited scalability	Limited scalability based on device capacity
Storage Capacity	Scalable and distributed storage systems	Limited storage capacity
Data Processing	Cloud infrastructure handles data processing	Local processing within edge or gateway devices
Real-time Analytics	Real-time or near real-time analytics	Batch processing or delayed analysis
Connectivity	Relies on internet connectivity	Various connectivity options (e.g., LAN, WAN)
Device Management	Centralized device management capabilities	Manual device management and updates
Flexibility and Interoperability	Wide range of APIs and tools for integration	Limited device compatibility and integration
Cost Structure	Pay-as-you-go or subscription-based pricing	Upfront investments in hardware and maintenance
Security	Robust security measures and compliance	Limited security measures
Analytics and Insights	Advanced analytics and machine learning	Basic data analysis or manual interpretation
Maintenance and Updates	Cloud service providers handle maintenance	Dedicated resources required for maintenance

9. IOT BASED ON CLOUD AND BIG DATA COMPONENTS

Our analysis reveals that a typical IoT system based on cloud computing and big data comprises the following essential components:

- IoT Devices: These are the physical devices such as sensors, actuators, and smart devices that collect and transmit data. They are responsible for capturing and sending data to the cloud.
- Cloud Infrastructure: This component forms the system’s foundational framework, encompassing cloud computing resources and services provided by major cloud service providers like AWS, Microsoft Azure, or Google Cloud Platform (GCP). It includes servers, storage, networking, and virtualization technologies.
- Data Storage Layer: It is a critical component that involves the storage of collected data, offering diverse storage technologies such as databases (both relational and NoSQL), data lakes, data warehouses, or distributed file systems. The choice of storage technology depends on specific IoT application requirements and data processing needs.
- Data Processing Layer: Performs a spectrum of data processing tasks on the collected data, including processing, transformation, aggregation, and analysis to derive valuable insights. Technologies like Apache Spark, Apache Hadoop, or cloud-based data processing services are commonly utilized in this layer.
- Analytics and Machine Learning: Encompasses advanced analytics techniques and machine learning algorithms to gain insights, identify patterns, and make predictions based on processed data. This component may feature tools for data visualization, dashboards, anomaly detection, predictive modeling, and recommendation systems.
- Security and Privacy: This component ensures the security and privacy of IoT data and the overall system by implementing measures such as authentication, access control, encryption, data anonymization, and compliance with data protection regulations.
- Integration and APIs: This component enables integration with external systems and services, allowing the exchange of data and interoperability with other applications. Application programming interfaces (APIs) facilitate communication and data sharing between different components or with external systems.
- Monitoring and Management: This component involves continuous monitoring of the system’s performance, health, and availability. It encompasses tools for real-time monitoring, logging, error handling, and alerting. Additionally, management tools enable configuration, scaling, and maintenance of system components.

These integrated components collaboratively form a scalable, flexible, and efficient cloud IoT big data system capable of managing the substantial volumes of data generated by IoT devices and extracting valuable insights. Fig. 5 illustrates the composition of IoT based on cloud and big data components.

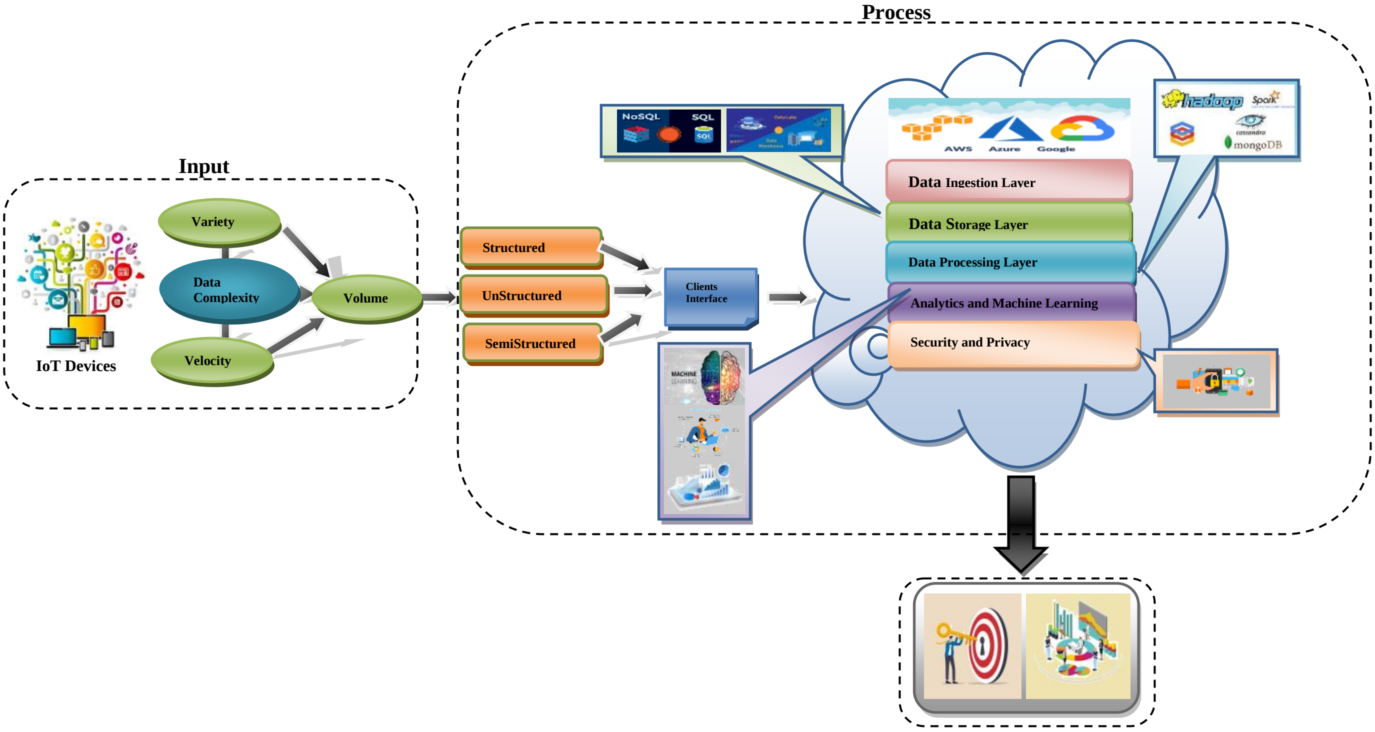

FIGURE 5. - **Components of IoT based on cloud and big data**

10. CHALLENGES AND OPEN ISSUES

The intersection of IoT, cloud computing, and big data offers exciting terrain for pioneering applications and groundbreaking solutions. Nevertheless, it introduces a unique set of hurdles and unresolved matters that necessitate attention for seamless execution. While the convergence of any two of these promising technologies has its intricacies and predicaments, Table 3 provides a succinct summary of exemplar challenges within each pairing.

Table 3. - Challenges between any pair of Internet of Things, cloud, and big data

Technology Pair	Open Problems and Challenges
IoT + Cloud	Scalability of cloud infrastructure for massive IoT data. Real-time data processing with low latency. Handling frequent connectivity loss between devices and cloud.
IoT + Big Data	Managing variety and velocity of streaming IoT data. Real-time analytics on heterogeneous IoT data. Extracting value from low-quality IoT data.
Cloud + Big Data	Resource optimization for analytics workloads. Unifying analytics across operational, IoT, and enterprise data. Ensuring data security and privacy across the cloud.

As mentioned in Table 3, it is important to note that there are ongoing advances in architectures, algorithms, and mechanisms that are needed to address issues as the technologies converge.

In this discussion, we delve into the enduring key challenges and open issues that persist within IoT integrated with cloud and big data platforms, as follows:

1. Scalability and Resource Management: Accommodating the expanding scale of IoT deployments, which includes managing soaring data volumes, device connections, and computational resources, remains a formidable challenge. The ability to efficiently manage resources and adapt dynamically to shifting demands is pivotal [31], [25].
2. Data Security and Privacy: Safeguarding the confidentiality, integrity, and privacy of IoT-generated data is an ongoing priority. This entails securing data during transmission, storage, and access, including robust encryption. Mitigating emerging threats and vulnerabilities is equally vital [5].
3. Interoperability: Within the vast IoT landscape, comprising diverse devices, protocols, and standards, achieving seamless interoperability and harmonious communication among these myriad components remains a persistent challenge [30].
4. Real-Time Processing: Many IoT applications hinge on real-time data processing for timely decision-making. Mitigating latency and ensuring responsive data processing pipelines are focal points of concern [8], [30].
5. Data Quality and Reliability: IoT data frequently arrives in noisy and incomplete forms. Maintaining data quality, reliability, and consistency to enable accurate analytics and decision support is an ongoing challenge [25], [7].
6. Edge Computing: The quandary of where to process data—whether at the edge or in the cloud—remains an open question. Striking the right balance between edge and cloud computing for specific use cases remains an elusive goal.
7. Energy Efficiency: Battery-powered IoT devices, often deployed in energy-constrained environments, demand enhanced energy efficiency without compromising functionality, an enduring challenge [7].
8. Standardization: The absence of universally accepted standards impedes the seamless integration and scalability of IoT, cloud, and big data systems. Ongoing efforts to standardize face formidable hurdles [36].
9. Regulatory Compliance: Meeting diverse data protection and privacy regulations, such as GDPR, while deploying IoT solutions is a multifaceted and cost-intensive task [22].
10. Data Governance: Developing robust data governance frameworks for IoT-generated data, encompassing aspects of data ownership, access, and lifecycle management, presents a formidable challenge [37].
11. Cost Management: As IoT deployments expand, the cost of utilizing cloud resources for data storage and processing can become prohibitive. Devising cost-effective strategies for resource allocation and management is essential [22], [17].
12. Ethical Considerations: Ethical quandaries arise from the collection and utilization of IoT data. Addressing issues related to data collection, usage, ownership, and transparency is an evolving challenge [34], [7].
13. Complexity of Analytics: Implementing advanced analytics and machine learning on extensive IoT datasets can strain computational resources. Streamlining analytics pipelines for efficiency, scalability, and real-time processing is an ongoing concern [30].
14. Edge AI and Machine Learning: Deploying AI and machine learning algorithms at the edge for immediate decision-making is an emerging frontier, marked by challenges related to model deployment, updates, and inference efficiency [36], [3].
15. Data Integration and Fusion: Integrating heterogeneous data sources, encompassing structured and unstructured data across diverse domains, for comprehensive analysis poses a complex challenge [25], [34].
16. Data Lifecycles: Managing the full lifecycle of IoT data, including retention, archival, and secure disposal, while aligning with stringent data protection regulations, presents intricate logistical challenges [19]. The emergence of quantum computing brings potential threats to established encryption methods, necessitating the development of quantum-resistant security measures [36].

In conclusion, the integration of IoT, cloud computing, and big data offers tremendous opportunities yet simultaneously presents a multitude of challenges and open issues. Researchers, industry stakeholders, and policymakers are engaged in collaborative efforts to address these intricacies and foster innovation in this dynamic and rapidly evolving realm.

11. CONCLUSION

The synergistic fusion of the IoT, cloud computing, and big data offers remarkable potential for transformation across various domains. As detailed in this paper, IoT continuously generates vast real-time data from interconnected devices and sensors. By harnessing the virtually limitless storage and computational resources of the cloud, this influx of IoT data can be efficiently processed. Subsequently, big data analytics unlocks invaluable insights through techniques such as machine learning, predictive modeling, and data visualization.

However, as highlighted in Sections 7 and 8, this convergence also ushers in notable challenges concerning scalability, security, privacy, real-time processing, and more. Strategies such as robust encryption, access control, data anonymization, resource optimization, edge computing, and standards development can effectively address these hurdles.

The analysis provided in this paper, including the comparative evaluation in Sections 5 and 6, enlightens us on how the synergistic integration of IoT, cloud, and big data can lead to transformative progress across industries. For instance, within healthcare, this integration can facilitate remote patient monitoring, personalized treatment, and preventive care by capitalizing on real-time data and analytics. In smart cities, it can streamline traffic flows, bolster public safety, and reduce energy consumption. The transformational impact also extends to sectors like manufacturing, transportation, and agriculture.

Fully realizing the potential of this technology convergence requires the collaborative efforts of researchers, businesses, policymakers, and other stakeholders. With a diligent focus on scalability, interoperability, security, privacy, and other vital issues, the synergistic integration of IoT, cloud computing, and big data can unlock a brighter future across industrial and societal domains.

Funding

None

ACKNOWLEDGEMENT

The authors are very grateful to the University of Mosul/College of Computer Science and Mathematics for the facilities they provided, which helped to improve the quality of this work.

CONFLICT OF INTERESTS

The authors declare no conflict of interest.

REFERENCES

[1] S. Khare and M. Totaro, “Big data in IoT,” in 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, Jul. 2019, pp. 1–7 [Online]. doi: 10.1109/ICCCNT45670.2019.8944495.
[2] H. Cai, B. Xu, L. Jiang, and A. V. Vasilakos, “IoT-based big data storage systems in cloud computing: Perspectives and challenges,” IEEE Internet Things J., vol. 4, no. 1, pp. 75–87, Feb. 2017. doi: 10.1109/JIOT.2016.2619369.
[3] Y. Chen, “IoT, cloud, big data and AI in interdisciplinary domains,” Simul. Model. Pract. Theory, vol. 102, p. 102070, Jul. 2020. doi: 10.1016/j.simpat.2020.102070.
[4] A. Hadjadj and K. HALIMI, “COVID-19 patients’ health monitoring system using fuzzy ontology and internet of things,” Iraqi J. Comput. Sci. Math., vol. 4, no. 1, pp. 191–203, Jan. 2023. doi: 10.52866/ijcsm.2023.01.01.0016.
[5] A. Malik and H. Om, “Cloud computing and internet of things integration: Architecture, applications, issues, and challenges,” in Sustainable Cloud and Energy Services, Cham: Springer International Publishing, 2018, pp. 1–24. [Online]. Available: doi: 10.1007/978-3-319-62238-5_1.
[6] H. Kumar, P. J. Soh, and M. A. Ismail, “Big data streaming platforms: A review,” Iraqi J. Comput. Sci. Math., vol. 3, no. 2, pp. 95–100, Apr. 2022. doi: 10.52866/ijcsm.2022.02.01.010.
[7] M. Talebkhah, A. Sali, M. Marjani, M. Gordan, S. J. Hashim, and F. Z. Rokhani, “IoT and big data applications in smart cities: Recent advances, challenges, and critical issues,” IEEE Access, vol. 9, pp. 55465–55484, 2021. doi: 10.1109/ACCESS.2021.3070905.
[8] M. Azeem, A. Haleem, S. Bahl, M. Javaid, R. Suman, and D. Nandan, “Big data applications to take up major challenges across manufacturing industries: A brief review,” Mater. Today Proc., vol. 49, pp. 339–348, 2022. doi: 10.1016/j.matpr.2021.02.147.
[9] A. H. Ali, M. Z. Abdullah, S. N. Abdul-wahab, and M. Alsajri, “A brief review of big data analytics based on machine learning,” Iraqi J. Comput. Sci. Math., vol. 1, no. 2, pp. 13–15, Jul. 2020. doi: 10.52866/ijcsm.2020.01.01.002.
[10] Z. S. Al-Talib and saad fawzi Al-Azzawi, “A new simple 6D hyperchaotic system with hyperbolic equilibrium and its electronic circuit,” Iraqi J. Comput. Sci. Math., vol. 4, no. 1, pp. 155–166, Jan. 2023. doi: 10.52866/ijcsm.2023.01.01.0013.
[11] M. Nazari Jahantigh, A. Masoud Rahmani, N. Jafari Navimirour, and A. Rezaee, “Integration of internet of things and cloud computing: A systematic survey,” IET Commun., vol. 14, no. 2, pp. 165–176, Jan. 2020. doi: 10.1049/iet-com.2019.0537.
[12] H. Chegini, R. K. Naha, A. Mahanti, and P. Thulasiraman, “Process automation in an IoT–Fog–Cloud ecosystem: A survey and taxonomy,” IoT, vol. 2, no. 1, pp. 92–118, Feb. 2021. doi: 10.3390/iot2010006.
[13] S. Selvaraj and S. Sundaravaradhan, “Challenges and opportunities in IoT healthcare systems: A systematic review,” SN Appl. Sci., vol. 2, no. 1, p. 139, Jan. 2020. doi: 10.1007/s42452-019-1925-y.
[14] F. Amin, R. Abbasi, A. Mateen, M. Ali Abid, and S. Khan, “A step toward next-generation advancements in the internet of things technologies,” Sensors, vol. 22, no. 20, p. 8072, Oct. 2022. doi: 10.3390/s22208072.
[15] O. Ali, M. K. Ishak, M. K. L. Bhatti, I. Khan, and K.-I. Kim, “A comprehensive review of internet of things: Technology stack, middlewares, and fog/edge computing interface,” Sensors, vol. 22, no. 3, p. 995, Jan. 2022. doi: 10.3390/s22030995.
[16] D. Serpanos and M. Wolf, Internet-of-Things (IoT) Systems. Cham: Springer International Publishing, 2018. [Online]. Available: doi: 10.1007/978-3-319-69715-4.
[17] I. Abrar, S. N. Pottoo, F. S. Masoodi, and A. Bamhdi, “On IoT and its integration with cloud computing:

Challenges and open issues," in Integration and Implementation of the Internet of Things Through Cloud Computing, 2021, pp. 37–64.
[18] N. Semwal, M. Mukherjee, C. Raj, and W. Arif, "An IoT based smart e-health care system," J. Inf. Optim. Sci., vol. 40, no. 8, pp. 1787–1800, Nov. 2019. doi: 10.1080/02522667.2019.1703269.
[19] S. Bharati, P. Podder, M. R. H. Mondal, and P. K. Paul, "Applications and challenges of cloud integrated IoMT," in Cognitive Internet of Medical Things for Smart Healthcare, Cham: Springer, 2021, pp. 67–85.
[20] G. Verma and S. Prakash, "Internet of things for healthcare: Research challenges and future prospects," in Advances in Communication and Computational Technology, Singapore: Springer, 2021, pp. 1055–1067.
[21] T. N. Nguyen, S. Lee, H. Nguyen-Xuan, and J. Lee, "A novel analysis-prediction approach for geometrically nonlinear problems using group method of data handling," Comput. Methods Appl. Mech. Eng., vol. 354, pp. 506–526, Sep. 2019. doi: 10.1016/J.CMA.2019.05.052.
[22] C. Doukas and I. Maglogiannis, "Bringing IoT and cloud computing towards pervasive healthcare," in 2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 2012, pp. 922–926. doi: 10.1109/IMIS.2012.26.
[23] G. Martinovic and B. Zoric, "E-health framework based on autonomic cloud computing," in 2012 Second International Conference on Cloud and Green Computing, IEEE, Nov. 2012, pp. 214–218. doi: 10.1109/CGC.2012.36.
[24] L. M. Dang, M. J. Piran, D. Han, K. Min, and H. Moon, "A survey on internet of things and cloud computing for healthcare," Electronics, vol. 8, no. 7, p. 768, Jul. 2019. doi: 10.3390/electronics8070768.
[25] D. B. Abdullah and Z. N. Al-Kateeb, "Prospects and challenges of the cloud of things in telemedicine," in 2022 8th International Conference on Contemporary Information Technology and Mathematics (ICCITM), IEEE, Aug. 2022, pp. 1–7. doi: 10.1109/ICCITM56309.2022.10031829.
[26] R. Kitchin and G. McArdle, "What makes big data, big data? Exploring the ontological characteristics of 26 datasets," Big Data Soc., vol. 3, no. 1, p. 205395171663113, Jun. 2016. doi: 10.1177/2053951716631130.
[27] S. Gupta and R. Godavarti, "IoT data management using cloud computing and big data technologies," Int. J. Softw. Innov., vol. 8, no. 4, pp. 50–58, Oct. 2020. doi: 10.4018/IJSI.2020100104.
[28] Y. Demchenko, C. de Laat, and P. Membrey, "Defining architecture components of the big data ecosystem," in 2014 International Conference on Collaboration Technologies and Systems (CTS), IEEE, May 2014, pp. 104–112. doi: 10.1109/CTS.2014.6867550.
[29] Buyya R., Calheiros R., and Dastjerdi A., Big Data: Principles and Paradigms. Elsevier, 2016, Big data: principles and paradigms. Elsevier, 2016.
[30] T. Alam, "Cloud-based IoT applications and their roles in smart cities," Smart Cities, vol. 4, no. 3, pp. 1196–1219, Sep. 2021. doi: 10.3390/smartcities4030064.
[31] E. Akanksha, A. Debnath, and B. Dey, "Extensive review of cloud based internet of things architecture and current trends," in 2021 6th International Conference on Inventive Computation Technologies (ICICT), IEEE, Jan. 2021, pp. 1–9. doi: 10.1109/ICICT50816.2021.9358598.
[32] Z. Sun, K. Strang, and R. Li, "Big data with ten big characteristics," in Proceedings of the 2nd International Conference on Big Data Research, New York, NY, USA: ACM, Oct. 2018, pp. 56–61 [Online]. doi: 10.1145/3291801.3291822.
[33] B. Unhelkar and V. Trivikram Rao, "Key considerations in optimizing the deployment of big data analytics-as-a-service utilizing cloud architecture and machine learning," 2020, pp. 818–832 [Online]. doi: 10.1007/978-3-030-30577-2_73.
[34] K. Kanagaraj and S. Geetha, "Data analytics framework based on cloud environment," in Integration of Cloud Computing with Internet of Things, Wiley, 2021, pp. 251–275. doi: 10.1002/9781119769323.ch15.
[35] C. Verma and R. Pandey, "Analytics-as-a-Service (AaaS)," in Web Services. IGI Global, 2019, pp. 1262–1281 [Online]. Available: doi: 10.4018/978-1-5225-7501-6.ch065.
[36] S. Krishnamoorthy, A. Dua, and S. Gupta, "Role of emerging technologies in future IoT-driven Healthcare 4.0 technologies: A survey, current challenges and future directions," J. Ambient Intell. Humaniz. Comput., vol. 14, no. 1, pp. 361–407, Jan. 2023. doi: 10.1007/s12652-021-03302-w.
[37] Y. Himeur et al., "AI-big data analytics for building automation and management systems: A survey, actual challenges and future perspectives," Artif. Intell. Rev., vol. 56, no. 6, pp. 4929–5021, Jun. 2023. doi: 10.1007/s10462-022-10286-2.