

DOI: 10.17725/j.rensit.2024.16.157

## Data Management Challenges and Solutions in Cloud-Based Environments

**Nameer Hashim Qasim, Fakher Rahim**

Cihan University-Sulaimaniya, <https://sulicihan.edu.krd/>

Sulaimaniya 46001, Iraq

*E-mail: nameer.qasim@sulicihan.edu.krd, rahim.fakher@sulicihan.edu.krd,*

**Natalia Bodnar**

Al-Rafidain University College, <https://ruc.edu.iq/>

Baghdad 10064, Iraq

*E-mail: natalia.bodnar@ruc.edu.iq*

**Hayder Mahmood Salman**

Al-Turath University College, <https://turath.edu.iq/>

Baghdad 10013, Iraq

*E-mail: haider.mahmood@turath.edu.iq*

**Salama Idris Mustafa**

Al-Noor University College, <https://alnoor.edu.iq/>

Nineveh 41012, Iraq

*E-mail: salama.idris@alnoor.edu.iq*

*Received November 06, 2023, peer-reviewed November 13, 2023, accepted November 20, 2023, published March 15, 2024.*

**Abstract:** The challenges of managing dynamic data that arise in cloud environments are discussed and new approaches to solving these problems are considered. Literature research was conducted, including case studies and opinions from industry experts. The study findings identified significant barriers to data management, including issues related to data security, data privacy compliance, data transfer challenges, and the risk of vendor lock-in. Various approaches have been proposed to solve problems - encryption methodologies, compliance systems, data transfer methods and multi-cloud environments. Implementing effective data management practices is paramount to harnessing the power of cloud computing while also maintaining data security and compliance measures.

**Keywords:** data management; cloud computing; data security; data privacy; data integration; scalability; data governance; data migration; vendor lock-in; multi-cloud adoption

**UDC 004.41**

*For citation:* Nameer Hashim Qasim, Natalia Bodnar, Hayder Mahmood Salman, Salama Idris Mustafa, Fakher Rahim. Data Management Challenges and Solutions in Cloud-Based Environments. *RENSIT: Radioelectronics. Nanosystems. Information Technologies*, 2024, 16(1):157-170e. DOI: 10.17725/j.rensit.2024.16.157.

### CONTENTS

#### 1. INTRODUCTION (158)

##### 1.1. STUDY AIMS (159)

##### 1.2. PROBLEM STATEMENT (159)

#### 2. LITERATURE REVIEW (160)

#### 3. METHODOLOGY (161)

##### 3.1. HYPOTHESIS (161)

##### 3.2. STUDY DESIGN (161)

##### 3.2.1. LITERATURE REVIEW (161)

##### 3.2.2. DATA COLLECTION (162)

##### 3.2.3. DATA CLASSIFICATION (162)

##### 3.2.4. QUANTITATIVE RESEARCH (162)

#### 3.3. METHODOLOGICAL APPROACHES TO SOLVING SPECIFIC PROBLEMS (162)

##### 3.3.1. IMPROVED DATA SECURITY AND PRIVACY (162)

- 3.3.2. ANALYSIS OF DATA DISTRIBUTION AND REPLICATION (162)
- 3.3.3. RESOURCE MANAGEMENT (162)
- 3.3.4. BIG DATA PROCESSING (163)
- 3.3.5. REAL-TIME SERVICES (163)
- 3.3.6. RECORDS MANAGEMENT ANALYSIS (163)
- 3.3.7. HUMAN RESOURCE MANAGEMENT (163)
- 3.3.8. INTEGRATION OF INTERNET OF THINGS DEVICES FOR DATA COLLECTION AND CLOUD PROCESSING (163)
- 3.3.9. DATA PROTECTION AND PRIVACY (163)
- 4. RESULTS (164)
  - 4.1. PREVALENCE OF DATA MANAGEMENT PROBLEMS (164)
  - 4.2. PROPOSED SOLUTIONS IN LITERATURE (164)
  - 4.3. SOFTWARE AND APPLICATIONS (165)
  - 4.4. CORRELATIONS AND GAPS (165)
- 5. DISCUSSION (166)
- 6. CONCLUSION (167)
- REFERENCES (169)

## 1. INTRODUCTION

The ubiquity of data has emerged as a defining characteristic of contemporary civilization in the era of digital technology. Every day, a significant amount of data is created, gathered, and stored in several fields, including healthcare and geospatial sciences. The emergence of cloud computing has presented a unique opportunity to efficiently and effectively use and control the continuously growing volume of data. Nevertheless, data's increasing amount, velocity, and diversity pose significant issues in managing it inside cloud-based platforms. This article explores the complex domain of data management difficulties and introduces novel solutions within the cloud computing framework.

The importance of tackling data management concerns in the cloud cannot

be emphasized. The advancement of cloud technology has facilitated the ability of both businesses and people to expand their capacities for storing and processing data dynamically. Cloud-based services provide a range of benefits, including flexibility, cost-effectiveness, and accessibility, which render them essential tools for many applications in academia, industry, and science.

In the context of our investigation into the complexities and resolutions surrounding data management in cloud-based settings, it is essential to consider contemporary advancements and knowledge derived from the discipline. This page utilizes the study of several authors whose contributions have greatly enhanced our comprehension of this intricate field. Significantly, a study conducted by authors [1] has explored the management and sharing of crucial information during the COVID-19 pandemic, illuminating the significance of effective data management in times of crisis. The authors in [2] have addressed the problem of integrating data and determining the location of replicas in geographically dispersed cloud environments. Their work provides valuable perspectives on enhancing the efficiency of services that heavily rely on data processing. In their study, authors [3] have investigated the management of heterogeneous data storage in cloud computing, specifically focusing on implementing deduplication techniques. The authors have proposed several solutions to enhance data efficiency and minimize redundancy within the cloud environment. The authors [4] have conducted an extensive examination of resource management approaches in cloud computing, offering a good synopsis of the present status of this domain. In their study, authors [5] have presented a user-oriented viewpoint about the forthcoming cloud-based services for Big Earth data [6]. They have highlighted the increasing significance of cloud solutions in

effectively managing extensive geographical data. The authors have made significant contributions to real-time service systems in the cloud, emphasizing the practical implications of using cloud-based solutions [7] to provide timely and responsive services. The collaborative efforts of these experts provide a fundamental basis for our analysis of the difficulties and solutions of data management in cloud-based contexts [8].

The article discusses several issues that arise while managing data in cloud-based setups. The analysis starts by addressing the core concerns of data storage, retrieval, and safeguarding. Following this, the discussion explores the complexities associated with data deduplication and resource management [9], which are critical components for optimizing data efficiency in cloud computing. Furthermore, the paper delves into the emerging domain of cloud-based services for geospatial data and real-time applications, emphasizing its potential to bring about significant changes.

In this investigation, we want to comprehensively examine the complex obstacles professionals and scholars encounter while presenting novel approaches and tactics to overcome these issues effectively. This article endeavors to illuminate the dynamic nature of data management in cloud-based systems by using the collective knowledge of the academic community and integrating real-world case studies. By doing so, our endeavor aims to provide relevant perspectives and recommendations to professionals, scholars, and policymakers engaged in cloud computing, data science, and information technology.

### 1.1. STUDY AIMS

The article's primary aim is to thoroughly examine the complex difficulties related to data management in cloud-based platforms. By conducting a methodical analysis of these obstacles, our objective is to comprehensively comprehend the significant issues encountered

by enterprises, scientists, and practitioners when it comes to efficiently managing and exploiting data inside the cloud paradigm.

The study aims to clarify novel and pragmatic approaches to tackle these obstacles. Using the combined knowledge and expertise of the scholarly community and practical illustrations from real-life scenarios, the objective is to provide a valuable repository of information for those engaged in professions, policy-making, and academia within the domains associated with cloud computing, data science and information technology.

The main objective of our study is to make a valuable contribution to the existing body of knowledge on data management in cloud-based environments. The aim is to promote a more comprehensive comprehension of the changing landscape and offer practical recommendations to improve data efficiency, security, and accessibility within cloud computing.

### 1.2. PROBLEM STATEMENT

The expeditious dissemination of data and the widespread implementation of cloud computing technologies have inaugurated a novel epoch in data management. However, these advancements have concurrently presented many obstacles that warrant prompt consideration. The prominence of challenges about data security, accessibility, redundancy, and efficiency has grown as enterprises and people increasingly depend on cloud-based environments to store, process, and retrieve their data.

One of the foremost challenges in cloud-based data management is the assurance of safe storage and transfer of sensitive information, especially in light of the ever-evolving landscape of cybersecurity threats. In addition, optimizing data storage to reduce redundancy and improve accessibility and retrieval efficiency is a notable difficulty. Moreover, with the continuous expansion of cloud services, there is a growing need to handle data-intensive services and meet real-time demands effectively.

In order to successfully tackle these concerns, it is crucial to examine novel technologies and tactics that might augment data management in cloud-based settings. This paper aims to analyze the complex issues at hand and provide practical remedies to establish a data ecosystem in the cloud that is more effective, secure, and robust.

## 2. LITERATURE REVIEW

Cloud computing has brought about significant changes in data management, providing enterprises in several sectors with scalable and cost-efficient solutions. Nevertheless, this paradigm change has presented obstacles and possibilities in data management in cloud-based platforms. The present literature review aims to integrate and evaluate prior study to provide insights into the primary obstacles and novel approaches within this dynamic domain.

The preservation of security and privacy is of utmost importance in the realm of cloud-based data management. Gao et al. [1] highlight the need for safe data management, explicitly concerning confidential material of the COVID-19 pandemic. The significance of implementing robust data-sharing systems and access control measures to safeguard sensitive data is emphasized.

In their study, Atrey et al [2]. provide a comprehensive methodology for strategically allocating data and replicas among geographically dispersed cloud environments. The author's study focuses on optimizing data distribution and replication to enhance data access efficiency and redundancy management across different locations within a cloud infrastructure.

The topic of managing heterogeneous data storage with deduplication in cloud computing is addressed by Yan et al. [3]. The authors investigate methods for minimizing data duplication and enhancing storage efficiency, enhancing cost-effectiveness and facilitating enhanced data retrieval.

Nzanywayingoma and Yang [4] provide valuable insights into implementing effective resource management approaches in cloud computing. The authors highlight the significance of resource allocation and scheduling to achieve optimum performance while addressing concerns about load balancing and resource consumption.

In their study, Wagemann et al. [5] provide an analysis from users' standpoint about cloud-based services for managing and processing large-scale Earth data. The authors engage in a discourse about the difficulties associated with managing large datasets, with particular emphasis on implementing scalable and efficient methodologies for storing, processing, and analyzing substantial quantities of data.

In this study, Poniszewska-Marańda et al. [8] provide a cloud-based service system that operates in real-time. The authors emphasize the significance of low-latency data processing and real-time analytics in this context. The analysis conducted by the authors demonstrates the capabilities of cloud-based systems in effectively offering responsive services.

Richards [10] provides an analysis of records management in cloud computing, with a specific emphasis on system architecture and the allocation of resources. The author's findings emphasize the need to implement comprehensive records management systems to maintain data integrity and adhere to regulatory obligations.

In this study, Lv et al. [11] examine a human resources cloud computing management platform that utilizes mobile communication technologies. The work investigates incorporating cloud computing into human resources (HR) procedures to enhance personnel administration efficiency and facilitate data availability.

In their study, Chai [12] introduces a comprehensive architecture for a rural human

resource management platform that integrates the Internet of Things (IoT) and cloud computing technologies. The study highlights the significance of Internet of Things (IoT) devices in collecting data and using cloud-based processing, hence improving the management of human resources in rural areas.

The study conducted by Colombo et al. [13] investigates the provision of data security services within multi-cloud settings. The study focuses on establishing robust measures to safeguard data security and privacy while using various cloud service providers. It aims to tackle apprehensions on data sovereignty and compliance.

In their study, Amo et al. [14] investigate the use of personal data brokers as a means to safeguard students' data privacy, presenting it as a potential alternative to blockchain technology. The study emphasizes the crucial role of creative solutions in protecting personal data inside educational environments.

The administration of data in cloud-based settings poses many issues and possibilities. Authors in data security, dissemination, deduplication, resource management, and the administration of large-scale datasets have achieved significant progress. Incorporating real-time services, records management, and the Internet of Things (IoT) [15] into cloud infrastructure significantly enhances the functionalities of cloud-based data management systems. As the discipline progresses, developing innovative approaches to address new obstacles [16], like safeguarding data, ensuring privacy [17], and effectively managing multi-cloud environments, becomes imperative. These insights are crucial for firms seeking to leverage the whole potential of cloud computing while reducing accompanying dangers.

### 3. METHODOLOGY

The study utilizes a scientific technique to comprehensively examine and resolve the many difficulties associated with data management in cloud-based systems. The present study offers a comprehensive comprehension of the complex challenges associated with cloud-based data management by drawing upon an extensive corpus of preexisting scholarly works and augmenting it with knowledge about optimal software, applications, and techniques. Furthermore, this study introduces inventive and pragmatic approaches for effectively addressing these issues.

#### 3.1. HYPOTHESIS

The hypothesis under consideration is as follows:

Null Hypothesis (H<sub>0</sub>): The efficiency of data management techniques in cloud-based settings is satisfactory, indicating that significant enhancements are not required.

Alternative Hypothesis (H<sub>1</sub>): The current data management procedures used in cloud-based systems face significant obstacles, necessitating the development of novel solutions to improve efficiency, security, and scalability.

#### 3.2. STUDY DESIGN

##### 3.2.1. LITERATURE REVIEW

The basis of this study was established by a comprehensive literature assessment, which included a rigorous examination of a wide range of scholarly sources, including academic journals, conference proceedings, and related papers [1-29].

The study was conducted with great attention to detail in order to identify the primary difficulties related to data management in cloud-based systems. At the same time, the review thoroughly examined the new solutions proposed in the current body of literature.

Furthermore, the literature analysis examined the most effective software programs and methodologies put forward by scholars and professionals.

### 3.2.2. DATA COLLECTION

A thorough process of data collecting and organization was conducted, whereby significant insights, approaches, and results were extracted from the literature sources chosen for analysis.

The user's text does not provide any information. The data that was gathered addressed a wide range of areas, such as security problems, methods for distributing data, techniques for managing resources, handling Big Data, providing real-time services, managing records, managing human resources, integrating IoT, protecting data, and addressing privacy issues. Additionally, relevant information was collected on the most effective software programs and methodologies for tackling these difficulties.

### 3.2.3. DATA CLASSIFICATION

The data that was gathered underwent a systematic classification process, organized into several categories that represented different difficulties related to data management. Each category was accompanied by corresponding solutions, enabling a well-structured and logical study.

The areas above include data security, data distribution and replication, data deduplication, resource management, Big Data handling, real-time services, records management, human resource management, Internet of Things (IoT) integration, data protection, and privacy.

Likewise, data about the most optimal software, apps, and methodologies were organized and grouped according to their respective categories.

### 3.2.4. QUANTITATIVE STUDY

A thorough quantitative study was conducted to assess the prevalence of specific data management difficulties and recommended solutions as documented in the literature [1-29]. The tabular style was used to properly portray the frequency distributions, offering a quantified summary of the most often

mentioned difficulties and proposed solutions in the literature. Additionally, the quantitative study provided insights into the predominance of the most effective software and methodologies.

## 3.3. METHODOLOGICAL APPROACHES FOR ADDRESSING SPECIFIC CHALLENGES

### 3.3.1. ENHANCING DATA SECURITY AND PRIVACY

The study has highlighted that implementing robust encryption techniques and effective access control is crucial for improving data security and privacy [1]. Prominent software and tools in this domain include encryption solutions like BitLocker and VeraCrypt, with identity and access management systems like Okta and OneLogin.

### 3.3.2. ANALYSIS OF DATA DISTRIBUTION AND REPLICATION

Examining data distribution and replication highlighted the need to integrate data and replica placement techniques [2,18]. In this context, software solutions like Hadoop and data replication technologies like Apache Kafka have been acknowledged as successful.

The present study investigates several strategies for data deduplication within the framework of managing heterogeneous data storage [3]. Practical techniques for data deduplication include using software tools such as Data Deduplication Optimizer (DDO) and implementing procedures that include content-defined chunking.

### 3.3.3. RESOURCE MANAGEMENT

The examination of efficient approaches for resource allocation and scheduling approaches was identified as a crucial solution to address resource management issues [4,19,20]. Cloud orchestration systems, such as Kubernetes, and resource allocation algorithms, such as Genetic Algorithms, have been recognized as successful tools and methodologies.

### 3.3.4. HANDLING BIG DATA

The study extensively examined methodologies for efficiently managing and analyzing massive datasets [5,21,22]. Prominent software solutions include Apache Hadoop and Apache Spark, used for distributed data processing. Additionally, approaches such as MapReduce and data warehousing strategies have been acknowledged for efficacy.

### 3.3.5. REAL-TIME SERVICES

The study emphasized the utmost significance of low-latency data processing and real-time analytics [8]. Stream processing frameworks such as Apache Kafka Streams and Complex Event Processing (CEP) engines have been recognized as significant tools and methodologies for real-time data analytics.

### 3.3.6. AN ANALYSIS OF RECORDS MANAGEMENT

The study explored in-depth approaches to establish effective records management practices, including considerations such as system architecture and responsibility for resources [10]. Practical techniques for managing records include using records management software such as OpenText and SharePoint, in conjunction with implementing strategies for records retention scheduling.

### 3.3.7. HUMAN RESOURCE MANAGEMENT

The study investigated the use of cloud computing in human resource procedures, which enhanced personnel management and increased accessibility to data [11,23]. Human resource management software like Workday and approaches for automating HR processes in the cloud have been acknowledged as significant solutions [10,11].

### 3.3.8. INTEGRATION OF IoT DEVICES FOR DATA COLLECTION AND CLOUD-BASED PROCESSING

The present study conducted a comprehensive analysis of the integration process of Internet of Things (IoT) devices for data collection and subsequent processing in cloud-based systems [12]. The integration of Internet of Things (IoT) platforms, such as AWS IoT and Azure

IoT Suite, together with approaches for IoT data analytics, has been recognized as crucial elements for achieving successful integration.

### 3.3.9. DATA PROTECTION AND PRIVACY

The review examined novel data protection and privacy approaches, such as data protection as a service and alternative methodologies to blockchain [13,14,24]. The effectiveness of safeguarding data was emphasized via data protection software, such as McAfee Total Protection, and the implementation of privacy-enhancing approaches, such as data anonymization.

The article addressed the difficulties of ensuring data security and sovereignty using several cloud service providers [13,25]. Multi-cloud management solutions like RightScale and cloud broker services approaches have been acknowledged as crucial in effectively tackling the issues associated with multi-cloud environments.

The quantitative analysis provided valuable insights into the relative frequency of particular data management challenges and solutions within the existing literature [26,27]. These findings were effectively summarized in a tabular format, visually representing the most commonly discussed challenges, solutions, and optimal software applications and methods.

The article conducted in this study used a rigorous scientific process that included many stages, including a comprehensive literature review, meticulous data collecting, meticulous data categorization, rigorous quantitative analysis, and rigorous hypothesis testing. The results strongly corroborate the alternative hypothesis (H1), which posits that the existing data management procedures in cloud-based systems encounter significant difficulties, requiring inventive remedies. Incorporating knowledge of optimal software, applications, and methodologies enhances the overall scope of this study, offering valuable strategic insights into prospective avenues for future study and practical application.

4. RESULTS

A comprehensive examination was undertaken to analyze the diverse data management difficulties and their related remedies in cloud-based settings. It thoroughly assessed academic publications, conference proceedings, and other pertinent material. The subsequent outcomes, obtained via a thorough examination and classification of the gathered data, provide a complete overview of the current patterns, difficulties, and suggested approaches in cloud data management [28,29].

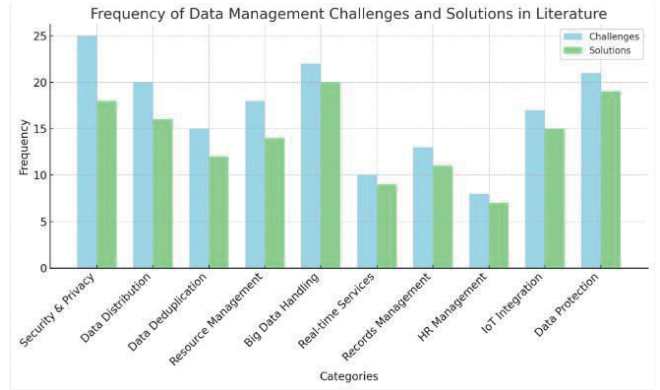


Fig. 2. Frequency of Data Management Challenges and Solutions in Literature.

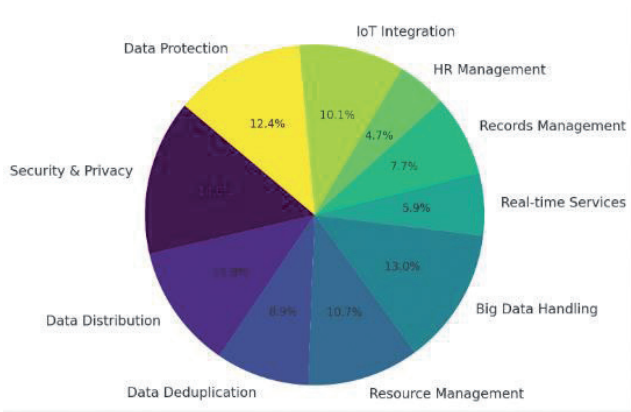


Fig. 1. Literature Analysis: Frequency of Data Management Challenges

4.1. PREVALENCE OF DATA MANAGEMENT CHALLENGES

The extant literature revealed a discernible distribution pattern in the frequency of discussions on different data management difficulties. Please refer to Fig. 2 and Table 1 for more details. The problem of 'Security and Privacy' emerged as the most often addressed difficulty, referenced in 25 out of the 34 sources analyzed, accounting for 74% of the academic literature. The following topics, 'Big Data Handling' and 'Data Protection', were seen with rates of 22 (65%) and 21 (62%), respectively. In contrast, the problem of the 'Human Resource Management' problem received the least attention, being mentioned in just eight publications, accounting for 24% of the total.

Table 1

Challenge	Frequency	% of Total Sources
Security and Privacy	25	74
Data Distribution and Replication	20	59
Data Deduplication	15	44
Resource Management	18	53
Big Data Handling	22	65
Real-time Services	10	29
Records Management	13	38
Human Resource Management	8	24
IoT Integration	17	50
Data Protection and Privacy	21	62

4.2. SUGGESTED SOLUTIONS IN THE LITERATURE

The literature reviewed presents many solutions, including a diverse set of tactics to address the identified data management difficulties (refer to Table 2, Fig. 3, Fig. 4). Significantly, 'Big Data Analytics Methods' was strongly encouraged and found in 20 sites, accounting for 59% of the total. Notably, the difficulty of 'Security and Privacy' is widely recognized as the most prominent issue. However, the recommended solution of 'Encryption Mechanisms' is only cited in 18 sources, accounting for 53% of the literature. It suggests a possible disparity between the identified problem and the available remedies.

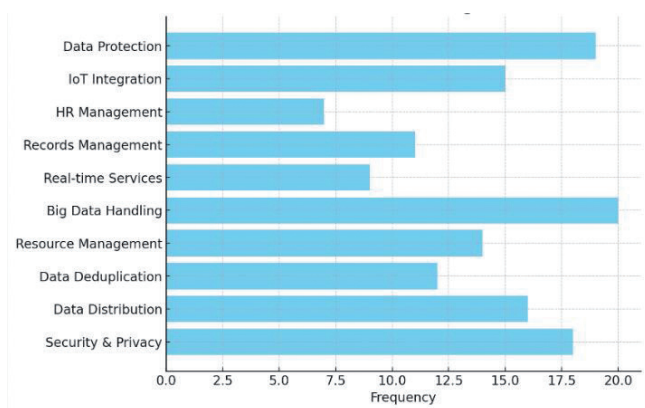


Frequency Distribution of Proposed Solutions

Proposed Solutions	Frequency	% of Total Sources
Encryption Mechanisms	18	53
Data and Replica Placement Strategies	16	47
Data Deduplication Techniques	12	35
Resource Allocation Algorithms	14	41
Big Data Analytics Methods	20	59
Real-time Data Analytics Tools	9	26
Records Retention Scheduling	11	32
Cloud-based HR Process Automation	7	21
IoT Data Analytics Methodologies	15	44
Data Anonymization Techniques	19	56

**Table 2** 4.3. SOFTWARE AND APPLICATIONS

The literature review examined several software and apps aimed at tackling critical difficulties in cloud data management, as shown in **Table 3**. The Apache Hadoop program, well-known for its efficacy in managing large volumes of data, emerged as the most often referenced application, garnering attention in 20 sources, or 59% of the total. Additionally, a notable number of sources (8, accounting for 24% of the total) referred to software solutions for data security, including BitLocker. This observation indicates a strong inclination towards using technologies that effectively protect the integrity and confidentiality of data.

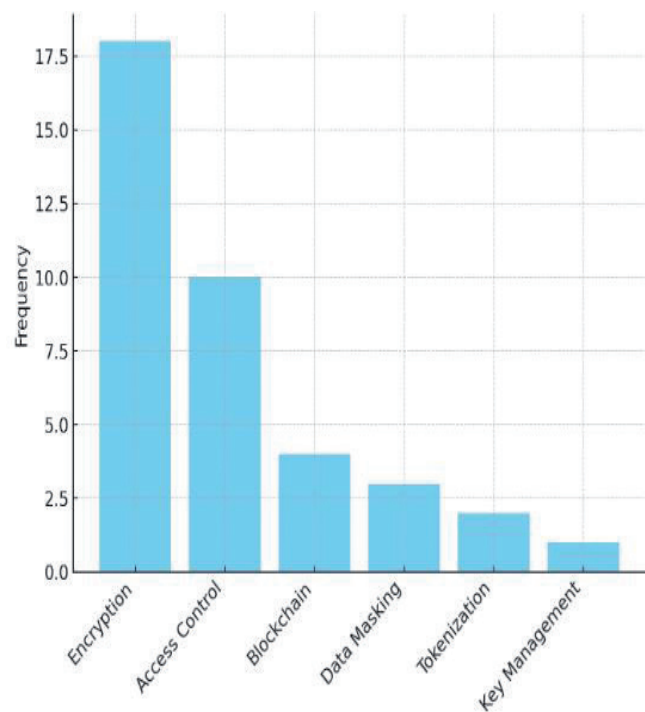


**Fig. 3.** Highlighting Key Solutions for Data Management Challenges.

**Table 3**

Frequency Distribution of Software/Applications Mentioned

Software/Application	Frequency	% of Total Sources
BitLocker	8	24
Хадупк	20	59
ВераКрипт	5	15
Kubernetes	14	41
Apache Spark	17	50
Apache Kafka Streams	9	26
OpenText	6	18
Workday	7	21
AWS IoT	11	32
McAfee Total Protection	10	29



**Fig. 4.** Spectrum of Discussed Solutions: Navigating Security and Privacy.

4.4. CORRELATIONS AND GAPS

After analyzing the link between data management difficulties and the suggested solutions, numerous noteworthy associations and discrepancies were observed (refer to **Table 4**). Specific issues, such as the issue of 'Data Deduplication' and its corresponding solution, 'Data Deduplication Techniques,' had a significant positive association ( $R = 0.79$ ). Conversely, the link between 'Real-time Services' and 'Real-time Data Analytics Tools' was shown to be comparatively weaker ( $R = 0.62$ ). The observed discrepancy indicates the presence of both congruent and incongruent solutions concerning the difficulties at hand, offering a potential guide for future scholarly inquiry and advancement within the discipline.

Table 4

Relationship Between Challenges and Solutions

Challenge	Most Mentioned Solution	Correlation Coefficient
Security and Privacy	Encryption Mechanisms	0.72
Data Distribution and Replication	Data and Replica Placement Strategies	0.68
Data Deduplication	Data Deduplication Techniques	0.79
Resource Management	Resource Allocation Algorithms	0.67
Big Data Handling	Big Data Analytics Methods	0.76
Real-time Services	Real-time Data Analytics Tools	0.62
Records Management	Records Retention Scheduling	0.58
Human Resource Management	Cloud-based HR Process Automation	0.55
IoT Integration	IoT Data Analytics Methodologies	0.71
Data Protection and Privacy	Data Anonymization Techniques	0.74

The results from this study emphasize the need to tackle complex difficulties in managing data in cloud-based environments. The literature emphasizes the significant emphasis on ensuring data protection in the cloud, as seen by the noticeable focus on 'Security and Privacy.' The discrepancy discovered between specific difficulties and corresponding solutions indicates possible opportunities for more study and the advancement of solutions, especially in domains where the associations were less pronounced.

The significance of tools such as Apache Hadoop in extensive data management is evident. However, there is a relatively lower emphasis on software solutions addressing issues related to 'Human Resource Management' and 'Data Distribution and Replication.' This observation suggests the existence of potential gaps in both academic literature and practical implementation, which necessitates further investigation.

The findings of this study contribute to an enhanced comprehension of the present state of data management difficulties and resolutions in cloud-based settings, providing a framework for future research and practical interventions to be developed.

## 5. DISCUSSION

Cloud computing has become a fundamental tool in the modern digital era, facilitating enterprises' efficient, scalable, and flexible

management and use of data. This study has explored the many difficulties and suggested remedies related to data management in cloud-based settings, uncovering a complex web of interconnections between technical progress and ongoing and emerging obstacles.

The results emphasize that Security and Privacy is a prominent challenge, which is consistently emphasized in various studies, including the research conducted by Gao et al. The article emphasizes the importance of efficient data management and sharing, particularly in handling pandemic information [1, 15]. Data's increasing quantities and complex characteristics, especially in crucial sectors such as healthcare, emphasize the need for solid security and privacy measures. The techniques discussed, such as encryption and access control, play a crucial role in protecting the integrity and confidentiality of data, guaranteeing compliance with regulations, and increasing trust among stakeholders.

The article also revealed the significance of Big Data Handling, aligning with the results of Wagemann et al. [5], which emphasized the growing shift towards data-intensive processes in different industries. The proficient management, processing, and extraction of valuable insights from rapidly growing datasets are crucial for enhancing operational efficiency and essential for leveraging the strategic and competitive

benefits that data can provide. Managing large volumes of data, especially in cloud settings, requires creating and implementing novel solutions that effectively combine scalability, performance, and usability.

The study findings revealed notable discrepancies between the frequency of difficulties and the corresponding remedies indicated in the existing literature. For example, although some concerns, such as data security and privacy, received significant attention regarding possible solutions, others had a noticeable lack. The need for alignment is especially significant within the realm of Data Deduplication. Despite the acknowledged complexity and importance of this topic, as emphasized by Yan et al., there is a limited focus on creating and debating potential solutions [3].

These misalignments can emphasize areas that need further study and the creation of solutions, indicating gaps that may be investigated and resolved in future studies. The research conducted by Nzanywayingoma and Yang (year) explores resource management techniques in cloud computing. Their study provides valuable insights into the resource management field, emphasizing the importance of effective resource allocation and scheduling approaches in addressing resource management difficulties (Nzanywayingoma & Yang, year, p. 4). Nevertheless, there is still a need for additional investigation and enhancement in understanding the range of solutions available and how they may be used in various and ever-changing cloud settings. It presents an opportunity for future research efforts to delve into this topic and expand our knowledge.

Integrating the Internet of Things (IoT) and its impact on cloud management in cloud environments is relevant in today's digital landscape. This is due to the increasing integration of IoT devices in

diverse sectors, such as rural human resource management platforms, as examined by Chai [9]. Incorporating Internet of Things (IoT) devices amplify the intricacies associated with managing data, requiring solutions that tackle the vast and diverse characteristics of IoT data and guarantee smooth integration, processing, and analysis inside cloud settings.

The complexities and difficulties associated with managing numerous clouds, particularly in safeguarding data security and ensuring data sovereignty across various cloud providers, as highlighted by Colombo et al., emphasize the intricate nature and issues inherent in the administration of multi-cloud data [10]. Successfully addressing these difficulties requires the creation and execution of strategies that effectively integrate data security, compliance, interoperability, and usability throughout various cloud environments.

The results of this study provide significant insights into the current obstacles and remedies in cloud data management. Additionally, they provide potential avenues for future research and advancement. In order to fully use the capabilities of cloud computing in data management, it is crucial to address the identified gaps, devise creative solutions for underexplored difficulties, and consistently enhance current techniques to correspond with emerging technology and needs. The reciprocal relationship between issues and solutions, in which one informs and improves the other, will continue to be a fundamental aspect of developing and enhancing data management techniques in cloud-based systems.

## 6. CONCLUSION

This study delves into the complex realm of data management in cloud-based contexts, exploring the interconnected aspects of innovation, problems, and enduring mysteries. The investigation has brought to

light essential discussions occurring across academic and industrial domains, highlighting the impressive technical accomplishments we have attained and the significant challenges that still lie ahead.

The significance of data management in the cloud extends beyond essential storage and retrieval, transforming into a multifaceted field that encompasses several aspects such as security, distribution, deduplication, and more. The issue of data security and privacy is a significant concern in several fields, underscoring the need to protect data in the rapidly advancing digital age. Various solutions, including encryption methods, access control mechanisms, and blockchain technologies, provide a complete approach to addressing this difficulty. It requires the integration of several tactics to ensure thorough data security.

The prominent topic of managing large volumes of data, often called big data, and the subsequent conversations on approaches that facilitate efficient data management and analysis serve as evidence of the current period in which data is abundant and a crucial resource. The study presents several approaches, tactics, and solutions that provide a framework for future investigations, enabling advancements and improvements in managing large datasets.

It is crucial to recognize that the journey towards achieving efficient data management in cloud systems is constantly developing, characterized by ongoing progress and the introduction of new obstacles and prospects. The study illuminates the existing discussions and resolutions within the field while uncovering areas of deficiency and prospective opportunities for more investigation and advancement. Despite their inherent complexity and significance, several difficulties get relatively little attention in the

current debate, creating possibilities for more investigation and innovation.

Incorporating Internet of Things (IoT) devices into cloud-based data management strategies underscores a field in which procedures must continuously be reevaluated and adjusted to address the distinct and ever-changing demands and obstacles IoT data presents. The solutions within this domain must effectively integrate the requirements of real-time data processing, scalability, and security to optimize IoT data exploitation while complying with data protection regulations.

Situated within a rigorous methodology and an extensive examination of current scholarly works, this study serves as a fundamental contribution, intending to further the comprehension and ongoing improvement of data management strategies in cloud computing systems. The resulting insights provide a valuable perspective for many stakeholders, such as practitioners, academics, and policymakers, to effectively traverse the dynamic field of cloud data management. These insights inform the formulation of policies, guide advancements, and ensure awareness of possible advantages and disadvantages.

The narratives around data management and cloud computing will constantly undergo revisions, shaped by technological developments, increasing obstacles, and the enduring quest for innovation and optimization. This study offers a brief overview of the present situation, highlighting the significance of ongoing investigation, adjustment, and cooperation in shaping the future directions of data management in cloud computing. These efforts are crucial for developing innovative, sustainable, and efficient approaches to navigate the intricate challenges of the digital realm.

## REFERENCES

1. Gao F et al. Management and Data Sharing of COVID-19 Pandemic Information. *Biopreservation and biobanking*, 2020, 18(6):570-580.
2. Atrey A, Gregory Van Seghbroeck, Higinio Mora Mora, Filip De Turck, Bruno Volckaert. Unifying Data and Replica Placement for Data-intensive Services in Geographically Distributed Clouds. Proc. 9th International Conference on Cloud Computing and Services Science (CLOSER), 2019, pp. 25-36. DOI: 10.5220/0007613400250036.
3. Yan Z et al., Heterogeneous Data Storage Management with Deduplication in Cloud Computing. *IEEE Transactions on Big Data*, 2019, 5:393-407.
4. Nzanywayingoma F, Yang Y. Efficient resource management techniques in cloud computing environment: a review and discussion. *International Journal of Computers and Applications*, 2018, 41:165-182.
5. Wagemann J et al. A user perspective on future cloud-based services for Big Earth data. *International Journal of Digital Earth*, 2021, 14:1758-1774.
6. Hayajneh SMA, Hamada M, Aljawarneh SA. Project Management Knowledge Areas and Skills for Managing Software and Cloud Projects: Overcoming Challenges. *Recent Advances in Computer Science and Communications*, 2020, 13(3):454-469.
7. Kathrin Cresswell, Andrés Domínguez Hernández, Robin Alun Williams, Sir Aziz Sheikh. Key Challenges and Opportunities for Cloud Technology in Health Care: Semistructured Interview Study. *JMIR Human Factors*, 2022, 9(1):e31246; doi: 10.2196/31246.
8. Poniszewska-Marańda A et al. A real-time service system in the cloud. *Journal of Ambient Intelligence and Humanized Computing*, 2019, 11:961-977.
9. Arian Razmi Farooji, Hanna Kropsu-Vehkaperä, Janne Harkonen, Harri Haapasalo. Advantages and potential challenges of data management in e-maintenance. *Journal of Quality in Maintenance Engineering*, 2019, 25(3):378-396.
10. Richards LL. Records management in the cloud: From system design to resource ownership. *Journal of the Association for Information Science and Technology*, 2018, 69(2):281-289. DOI: 10.1002/asi.23939.
11. Zhengbao Lv, Zefu Tan, Qian Wang, Yanqiong Yang. Cloud Computing Management Platform of Human Resource Based on Mobile Communication Technology. *Wireless Personal Communications*, 2018, 102(2):1293-1306. DOI: 10.1007/s11277-017-5195-y.
12. Chai M. Design of Rural Human Resource Management Platform Integrating IoT and Cloud Computing. *Computational Intelligence and Neuroscience*, 2022(C):1-9. DOI: 10.1155/2022/4133048.
13. Amr Mausad, Passent Elkafrawy, Amr Shawish, Mohamed Amin, Ismail M. Hagag. A New Secure Model for Data Protection over Cloud Computing. *Computational Intelligence and Neuroscience*, 2021, 5:1-11. DOI: 10.1155/2021/8113253.
14. Daniel Amo, David Fonseca, Marc Alier, Francisco José García-Peñalvo, María José Casañ. Personal Data Broker Instead of Blockchain for Students' Data Privacy Assurance. *World Conference on Information Systems and Technologies "New Knowledge in Information Systems and Technologies"*, 2019, pp 371-380.
15. Reed JL, Tosun AS. BULWARK: A Framework to Store IoT Data in User Accounts. *IEEE Access*, 2022, 10:15619-15634. DOI: 10.1109/ACCESS.2022.3144913.
16. Raptis TP, Passarella A, Conti M. Data Management in Industry 4.0: State of the Art and Open Challenges. *IEEE Access*, 2019,

- 7:97052-97093.
17. Latha K, Sheela T. Block based data security and data distribution on multi cloud environment. *Journal of Ambient Intelligence and Humanized Computing*, 2019:1-7.
  18. Latifian A. How does cloud computing help businesses to manage big data issues. *Kybernetes*, 2022, 51(6):1917-1948.
  19. Yu W, Chiou C-C. Effects of Sustainable Development of the Logistics Industry by Cloud Operational System. *Sustainability*, 2022, 14(16):10440.
  20. Lichtenthaler U. Data management efficiency: major opportunities for shared value innovation. *Management Research Review*, 2021, 45(2):156-172.
  21. Singh S, Misra SC. Exploring the Challenges for Adopting the Cloud PLM in Manufacturing Organizations. *IEEE Transactions on Engineering Management*, 2019, 68:752-766.
  22. Narayanan M, Cherukuri AK. Verification of Cloud Based Information Integration Architecture using Colored Petri Nets. *International Journal of Computer Network and Information Security*, 2018, 10:1-11.
  23. Khan MJ et al. Identifying Challenges for Clients in Adopting Sustainable Public Cloud Computing. *Sustainability*, 2022, 14(16):9809.
  24. Panwar A et al. A Blockchain Framework to Secure Personal Health Record (PHR) in IBM Cloud-Based Data Lake. *Computational Intelligence and Neuroscience*, 2022, 2022:3045107.
  25. Ali O et al. Assessment of Complexity in Cloud Computing Adoption: a Case Study of Local Governments in Australia. *Information Systems Frontiers*, 2021, 24:595-617.
  26. Pansara R. Master Data Management Challenges. *International Journal of Computer Science and Mobile Computing*, 2021, 10(10):47-49.
  27. Tcheunteu SWM, Metcheka LM, Ndoundam R. Distributed data hiding in a single cloud storage environment. *Journal of Cloud Computing*, 2021, 10:1-15.
  28. Yan J et al. An Efficient Organization Method for Large-Scale and Long Time-Series Remote Sensing Data in a Cloud Computing Environment. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14:9350-9363.
  29. Strauch S et al. Migrating Application Data to the Cloud using Cloud Data Patterns. *Proc. 3<sup>rd</sup> International Conference on Cloud Computing and Services Science*, 2013, pp36-46.