big data and cognitive computing

MDPI

Review

An Overview on the Challenges and Limitations Using Cloud Computing in Healthcare Corporations

Giuseppe Agapito [1,2,3,*] and Mario Cannataro [3,4]

[1] Department of Law, Economics and Social Sciences, University "Magna Græcia" of Catanzaro, 88100 Catanzaro, Italy
[2] "Cultura Romana del Diritto e Sistemi Giuridici Contemporanei" Research Center, University "Magna Græcia" of Catanzaro, 88100 Catanzaro, Italy
[3] Data Analytics Research Center, University "Magna Græcia" of Catanzaro, 88100 Catanzaro, Italy; cannataro@unicz.it
[4] Department of Medical and Surgical Sciences, University "Magna Græcia" of Catanzaro, 88100 Catanzaro, Italy
* Correspondence: agapito@unicz.it

Abstract: Technological advances in high throughput platforms for biological systems enable the cost-efficient production of massive amounts of data, leading life science to the Big Data era. The availability of Big Data provides new opportunities and challenges for data analysis. Cloud Computing is ideal for digging with Big Data in omics sciences because it makes data analysis, sharing, access, and storage effective and able to scale when the amount of data increases. However, Cloud Computing presents several issues regarding the security and privacy of data that are particularly important when analyzing patients' data, such as in personalized medicine. The objective of the present study is to highlight the challenges, security issues, and impediments that restrict the widespread adoption of Cloud Computing in healthcare corporations.

Keywords: cloud computing; big data; omics data; healthcare; artificial intelligence; cryptography; IoT; edge computing

Citation: Agapito, G.; Cannataro, M. An Overview on the Challenges and Limitations Using Cloud Computing in Healthcare Corporations. *Big Data Cogn. Comput.* **2023**, *7*, 68. https://doi.org/10.3390/bdcc7020068

Academic Editor: Carson K. Leung

Received: 13 March 2023
Revised: 27 March 2023
Accepted: 29 March 2023
Published: 6 April 2023

1. Introduction

The investigation of all living organisms and complex diseases, e.g., yeast, human, cancer, and Alzheimer's has highlighted the need for a new holistic vision to shed light on the multiple interactions among several biological players, such as genes, enzymes, and small molecules. In the reductionist approach [1], a single mutation or weakness is responsible for diseases and phenotype aberrancies. In contrast, the holistic approach [2] asserts that conditions and phenotype aberrancies are due to the intricate interactions network among several biological players.

The appearance of omics sciences [3] provides the approaches to consolidate the holistic idea mandatory for studying living organisms at all structural and functional levels, including humans. Omics includes the domains ending in -omics, such as proteomics, epigenomics, metabolomics, and microbiomics. In particular, the rapid advances in High-Throughput (HT) and Molecular Biology (MB) technologies make omics sciences a central part of medical research. The continuous technological advances in HT and MB have made it possible to comprehensively analyze a simple living organism's genome, e.g., a single bacteria, and a complex organism, e.g., humans, in a few hours or a few days [4,5]. The highest quality of HT and MB produces massive data per single experiment, transforming biology and genomics into data-driven sciences [6]. Only the practical analysis of this enormous amount of data will allow us to understand the complex aberrancies starting from the genome.

The transition of life sciences toward data-driven science provides researchers with new opportunities, making it possible to yield vast amounts of omics data in a cost-

and time-efficient manner. Simultaneously, acquiring, storing, distributing, analyzing, and interpreting these data is challenging [7]. The high data heterogeneity in terms of type and source requires technical improvements in many Information Technology (IT) domains, raising various privacy, security, storage, sharing, processing, and computing power issues. Hence, it is essential to develop specific algorithms and software tools for analyzing the different available types of omics data, such as protein sequences, single nucleotide polymorphisms (SNPs), and gene expressions, necessary for understanding the expression of genes and their regulation and the mutations in DNA underlying genetic diseases. A further contribution is the development of graphic interfaces that effectively display information from various data sources.

To meet these challenges, Cloud Computing and High-Performance Computing (HPC) architectures can significantly improve the speed of omics data investigation, analysis, reliability, and reproducibility.

Architectures based on multiprocessors, even multi-core, Graphics Processing Units (GPUs), and hybrids architectures, e.g., holding both GPUs and CPUs, make HPC architectures ideal for handling computations requiring significant amounts of computing power and memory. The strength of HPC systems is the extreme computational power obtained through parallel or distributed computing.

Parallel programming enables us to write code in order to take advantage of the multiple computational cores of modern CPUs. Parallel programming decomposes programs, e.g., the process, as several independent bunches, e.g., the threads allowing parallel and concurrent execution. Partitioning programs into smaller threads allows the exploitation of multiple cores within modern CPUs. Multiple cores on a single machine share memory. Hence, threads can be executed simultaneously using shared memory to synchronize and communicate with each other. A popular environment for threads is Java thread, while CUDA is a popular environment for exploiting the computational properties of Graphics Processing Units (GPU). Distributed computing uses network protocols such as TCP/IP, allowing applications to send and receive data to each other over the network by providing the services and protocols for exchanging data. Hence, a distributed application is built upon several layers. At the lower level, the network connects devices, allowing communication among them. At the higher level, services are defined on the network protocols. Finally, distributed applications run on top of these layers to perform tasks across the network. A popular library for distributed computing is Message Passing Interface (MPI) [8], which is available for many programming languages and architectures. Hence, parallel and distributed computing allows for solving complex problems in a short time by employing many computing resources simultaneously that would otherwise require a lot of time if performed sequentially.

Thus, programmers must explicitly develop parallel programs, e.g., in a global environment using a multi-threading paradigm or in a distributed environment through the Message-Passing Interface (MPI) standard [8], to exploit the computational power delivered from HPC systems. In addition, to ensure that HPC systems run at optimal performance, a suitable technical support service is required. All these constraints introduce additional expenses, e.g., purchase, maintenance, and development, making the HPC systems ideal for large IT research centers and limiting the spread in biological, medical, and genomics research centers. The limiting element for the significant employment of HPC is nowadays primarily computational. On the other hand, Cloud Computing [9,10] brings a new paradigm from the analogy with existing infrastructures, such as electricity or water. Consequently, the achievement of the results is guaranteed independently of where data or instruction sequences are stored or executed. When opening a tap or turning on a lamp, one does not wonder where the water or electricity comes from; the important thing is that these are made available. Similarly, when some commands or services need to be executed in the Cloud system, it does not matter who takes care of it; the overall system will have to deliver the correct results based on the user requests. Thus, Cloud Computing provides an on-demand system through the Internet. Therefore, it eliminates purchase,

maintenance, and development costs, making high-performance computation available even for small research centers. Cloud Computing is available in three fundamental models, such as IaaS (Infrastructure as a Service), PaaS (Platform as a Service), and SaaS (Software as a Service).

Cloud Computing could be the ideal tool for dealing with many steps of the bioinformatics analysis pipeline, from pre-processing, selection, aggregation, and analysis, including exploration and visualization.

To take full advantage of the considerable benefits of Cloud Computing, healthcare corporations must face several management, technology, security, and legal issues that affect its rapid adoption in healthcare. For example, storing confidential health information in third-party remote data storage raises serious problems related to the patient's sensitive information because patient data could be lost or misused.

Thus, the present study aims to highlight the challenges, security issues, and impediments that limit the adoption of cloud computing in healthcare corporations.

The rest of the manuscript is arranged as follows: Section 2 describes the principal service and deployment models of Cloud Computing, highlighting the main difference between them. Section 3 introduces some well-known Cloud services suitable for handling Big omics Data. Section 4 describes challenges, security issues, and impediments that are limiting the spread of Cloud Computing in healthcare corporations. Section 5 discusses some of the possible challenges and issues to address underlying the low adoption of Cloud services in healthcare. Section 6 provides some guidelines to follow, when dealing with Cloud Computing in healthcare. Finally, Section 7 concludes the manuscript.

2. Cloud Computing

Cloud infrastructures comprise the front end and back end. The front end refers to the end users' devices (e.g., pc, tablets, or smartphones), an Internet connection, and a web browser or similar application indispensable to accessing the Cloud Computing environment. Two different types of users can benefit from the front end: (i) the user of the final Cloud service; (ii) the developer and owner of the provided Cloud service. Through the front end, the provider ensures the final users that data on its hosts are always available through Internet connections. Simultaneously, developers can always have access to enhance and maintain their services by interacting with the Cloud system through terminals-scripts, RESTfull services [11], and even using traditional browsers. The back end includes the data center resources providing security, storage capacity, and computing power necessary to keep all the Cloud ecosystems available to the users.

2.1. Service Models of Cloud Computing

Cloud Computing includes different standardized service models, among which are the following:

- Software as a Service (SaaS) [12] allows the use of the provider's applications running on remote architectures. The applications are obtainable through client applications, such as a web browser or an Application Program Interface (API). Users cannot control or manage the beneath Cloud infrastructure components such as network, servers, operating systems, storage, or individual application capabilities, excluding determinate user-specific application configuration settings.
- Platform as a Service (PaaS) [13] enables users to develop in the Cloud environment the users' applications created using libraries, services, and APIs compatible with the Cloud provider. Users cannot directly manage or control the infrastructure beneath the Cloud, including network, servers, operating systems, or storage, but retain the deployed applications and particular configuration settings for the application-hosting domain.
- Infrastructure as a Service (IaaS) [14] facilitates the user in provision processing, storage, networks, and other essential computing resources where the user can deploy and run the software, including operating systems and apps. Users cannot manage or

control the beneath-Cloud infrastructure, whereas having control of the operating systems, storage, deployed applications, and limited control on some select networking components, e.g., host firewalls or bridges.

Over the years, in addition to the essential service models, new Cloud service models have been added, including the following:

- Business Process as a Service (BPaaS) [15] exploits the Cloud to automate and drive down the costs of business processes carried out by organizations.
- Data as a Service (DaaS) [16] offers Cloud-based Big Data cleaning, filtering, and enrichment schemes to produce data sets suitable for predictive or prescriptive analyses.
- Connectivity as a Service (CaaS) [17] provides Voice-Over-IP (VOIP), video-conferencing, and Instant Messaging (IM) functions as Cloud-based subscription services for commercial institutions.
- Identity as a Service (IDaaS) [18] provides Cloud-based centralized authentication and Single-Sign-On (SSO) services on heterogeneous or federated Cloud schemes.

A critical aspect of each Cloud services model is the Multi-Tenancy (MT). MT is the Cloud platforms' power to satisfy multiple user requests concurrently, providing the highest separation between run time environment and data. MT is achieved by virtualizing the applications' run time environment and/or operating system, allowing users' applications to run on different Virtual Machines (VM). MT differs from multi-user operations, where multiple users share the same application. Still, the user applications and run time data, also known as user context, are only logically separated, e.g., held in different files or directories on the same physical storage.

2.2. Deployment Models of Cloud Computing

Deployment models have been developed alongside cloud service models to support users' business workloads. Today, business applications and processes rely on a complex ecosystem of hardware and services, each with its prerequisites in terms of privacy, availability, and scalability. Over the last decade, Cloud Computing has been embraced to improve business processes, and its models have been extended to meet the challenges in various scientific areas, including healthcare.

The Cloud Computing deployment models include the following:

- Public Cloud infrastructure is ideal for organizations needing quick access to computing resources without significant capital expenditure. Public Cloud infrastructure allows organizations to purchase virtualized computing services through the Internet. Since Public Cloud services are furnished as pay-per-use, no initial investments are required because new resources can be purchased when needed. Public Cloud services are ideal for healthcare organizations that cannot afford an investment in particular hardware and maintenance.
- Private Cloud infrastructure is intended for exclusive use by a single organization. The Private Cloud lets organizations complete control over how data are shared and stored, an optimal solution if security is the primary concern, e.g., in the healthcare domain, ensuring compliance with any ethical regulations and protecting the subject's sensitive data. Additionally, the Private Cloud provides on-demand data availability, guaranteeing trustworthiness and support for mission-critical tasks.
- Hybrid Cloud infrastructure combines Public and Private Cloud infrastructures by allowing data and applications to be moved between them. Cloud infrastructures are unique entities linked by standardized or proprietary technologies, enabling the portability of data and applications. Hence, Hybrid Cloud provides a unique integrated environment combining locally Private and Public Cloud services. Healthcare organizations using Hybrid Cloud could enhance the standard of security. In this regard, data and services that do not affect sensitive information can be available through the Public Cloud. In contrast, sensitive information held in the Private Cloud are under the institution's absolute control.

- Multicloud infrastructure handles several Cloud services by different providers, including organizations' Private Cloud resources and private computational assets, to accomplish various requirements and demands in a single heterogeneous Cloud environment. Multicloud gives more flexibility regarding service and computational capabilities, improving performance and increasing resource availability and redundancy, letting organizations and final users to use all available resources efficiently.
- Federated Cloud infrastructure is a heterogeneous Cloud environment connecting diverse providers through a partnership mechanism, e.g., a standard policy to share, access, and control infrastructure and services. Federated Cloud commonly combines multiple Private and Public Clouds. Federation members remain independent in resource sharing and access control, comprising federated identity management. Thus, the Federated Cloud increases reliability and, simultaneously, the scaling up of resources.
- Intercloud is a general model of Cloud infrastructures that incorporates heterogeneous Clouds from various providers and typically includes non-cloud resources. Intercloud models may use the Federated Cloud standard as the basis for creating or implementing more specific but customized control and management functions.

To sum up, the Public Cloud is suitable for use cases in which it is necessary to scale up quickly, execute short-term jobs, and mitigate the request for computational resources. The Private Cloud is ideal for use cases in which it is mandatory to protect sensitive information, including patents, meet data compliance requirements, ensuring high availability. The Federated Cloud infrastructure enables application scalability and workload optimization requirements through a federation paradigm between Public and Private Clouds. Hybrid Cloud is ideal for combining Public and Private Cloud services on-site in a unique integrated architecture. Multicloud is ideal for using multiple Cloud services, even from different providers. Multicloud can also incorporate physical and virtual infrastructures in a single heterogeneous Cloud environment. Intercloud is ideal for implementing more specific but customized common control and management functions for creating a virtual Private Cloud with restricted access based on federated access.

Table 1 shows the advantages and disadvantage of Deployment models.

Table 1. The table summarises the advantages and disadvantages of Cloud Deployment Models. In the table DM are the initials of Deployment Models; CP refers to Computational Power; S indicates the Security; AS introduces the Applications Scalability; AP denotes the Applications Portability; ToJ refers to Type of Job; HS refers to Heterogeneous Service; C refers to the Costs; EU indicates the Exclusive Use; T is the Trustness; sj, cj, and gj are the initials of short, critical, and general job, finally, the $\surd$ indicate feature availability, while $\times$ indicates absence of the feature.

DM	CP	S	AS	AP	ToJ	HS	C	EU	T
Public	$\surd$	$\times$	$\surd$	$\times$	sj	$\times$	$\times$	$\times$	$\times$
Private	$\surd$	$\surd$	$\surd$	$\times$	cj	$\times$	$\surd$	$\surd$	$\surd$
Federate	$\times$	$\surd$	$\surd$	$\times$	cj	$\surd$	$\surd$	$\surd$	$\surd$
Hybrid	$\times$	$\times$	$\times$	$\surd$	gj	$\surd$	$\surd$	$\times$	$\times$
Multicloud	$\surd$	$\times$	$\surd$	$\times$	gj	$\surd$	$\surd$	$\times$	$\times$
Intercloud	$\times$	$\surd$	$\times$	$\times$	gj	$\surd$	$\surd$	$\surd$	$\surd$

3. Background

Healthcare organizations generate a vast range of data and information. Thanks to the progress of HT omics technologies, there has been an exponential growth of omics data, e.g., gene expressions, sequences alignment, and protein sequences, rendering classical computational approaches ineffective for handling these massive amounts of heterogeneous data. Consequently, omics sciences turned into Big Data science. Big Data in health and medical areas need infrastructures to improve data storage and management. Data sharing and security are critical in health and medical care since researchers need easy and extensive access to data for scientific analysis and sharing results. Cloud Computing solutions for healthcare organizations can contribute to making data analysis, sharing,

access, and storage effective through Cloud services able to scale when the amount of data increases. Thus, Cloud Computing services are a cost-effective solution for storing, accessing, analyzing, sharing, and protecting healthcare data and information.

The following is a list of well-known Cloud services models suitable for handling Big omics Data.

- Cloud BioLinux [19] provides a platform for developing bioinformatics infrastructures on the Cloud. Cloud BioLinux is a publicly accessible Virtual Machine (VM) to create on-demand frameworks for high-performance bioinformatics computing using Cloud architectures. Cloud BioLinux preconfigured command line and graphical software applications are available through the Amazon EC2 Cloud. Cloud BioLinux is distributed under the MIT Licence, including different Cloud BioLinux VMs, whereas source code and user guides are available at http://www.cloudbiolinux.org (accessed on 21 March 2023).
- Cloud4SNP [20] is a Cloud-based framework for the parallel preprocessing and statistical analysis of pharmacogenomics SNP DMET microarray data sets. Cloud4SNP extends the DMET-Analyzer [21] engine to be implemented as a Cloud Computing service through the Data Mining Cloud Framework [22]. Data Mining Cloud Framework is a software framework for creating and implementing knowledge discovery workflows on the Cloud [23]. Cloud4SNP performs massive statistical tests of SNPs relevance in case-control studies using the well-known Fisher test. Cloud4SNP exploits data parallelism and employs an optimized filtering technique to bypass the execution of ineffective Fisher tests by removing rows, e.g., probes with similar SNPs distributions.
- CloudBurst [24] is a parallel read-mapping algorithm optimized for mapping Next-Generation Sequence (NGS) data from several organisms, including homo sapiens, SNPs discovery, genotyping, and personal genomics. CloudBurst runs the short Read-Mapping Program (RMAP) linearly since running time decreases linearly with the number of reads mapped, reaching a linear speedup increasing the number of processors. These results are obtained by implementing Hadoop MapReduce [25] to parallelize execution using multiple computing nodes. In this way, CloudBurst improves performance by decreasing the running time to minutes for mapping millions of short reads to the human genome. CloudBurst is available as an open-source Java project for Amazon EC2 at https://sourceforge.net/projects/cloudburst-bio/ (accessed on 21 March 2023).
- CloudMan [26] is a Cloud manager that directs all of the steps required to create and control a complete data analysis environment on a Cloud infrastructure using a web browser. CloudMan provides an NGS analysis technique integrated with the Galaxy applications. CloudMan comes with a graphical interface to enable an easy access to Cloud Computing services. CloudMan is currently available for Amazon Web Services (AWS) Cloud infrastructure as part of the Galaxy Cloud [27] and CloudBioLinux [28].
- Crossbow [29] is a scalable, portable, and automatic Cloud service for identifying SNPs from high-coverage short-read resequencing data. Crossbow implements the MapReduce framework [25] distributed from Apache Hadoop. Alignment and variant calling in Crossbow are performed using the Bowtie [29] and SOAPsnp [30] software tools.
- Eoulsan [31] is a Cloud service implementing the Hadoop MapReduce approach devoted to HT sequencing RNA-seq data analysis. The Eoulsan differential analysis of transcript expression workflow comprises six steps: (i) quality control filtering; (ii) reads mapping; (iii) alignments filtering; (iv) transcript expression calculation. (v) normalization; (vi) detection of significant differential expression. Eoulsan is available as standalone, local cluster, or Cloud Computing on Amazon Elastic MapReduce (EMR).
- Eoulsan 2 [32] is the update of Eoulsan initially developed for analyzing RNA-seq data. Eoulsan 2 introduces the following updates to handling long-read RNA-seq and scRNA-seq data: (i) enhances the workflow manager; (ii) facilitates the development of new modules; (iii) expands its applications to long-read RNA-seq and scRNA-seq.

Eoulsan 2 is implemented in Java, available only for Linux systems, and distributed under the LGPL and CeCILL-C licenses at http://outils.genomique.biologie.ens.fr/eoulsan/ (accessed on 21 March 2023). The source code and sample workflows are available on GitHub https://github.com/GenomicParisCentre/eoulsan (accessed on 21 March 2023).

- HealtheDataLab [33] is a Cloud Computing platform for analyzing Electronic Medical Records (EMRs) data with computing capability for analyzing Big Data. HealtheDataLab enables the building of statistical and machine learning models flexibly through the use of Amazon Web Services (AWS), allows for scalability and high-performance computing system, and complaints with the Health Insurance Portability and Accountability Act (HIPAA) standard. HealtheDataLab is available upon request made directly to Cerner Corporation.
- iMage Cloud [34] allows the analysis of medical images integrated with EMRs, enabling the sharing of images, EMRs, and merged images via the Internet. iMage uses Hybrid Cloud to deliver more convenient and secure services, allowing high-performance image processing and virtual applications to be delivered securely, conveniently, and efficiently. iMage provides a graphical user interface with which it is possible to share images after being combined with EMRs.
- PeakRanger [35] is a software package that resolves closely spaced peaks obtained from Chromatin Immunoprecipitation (ChIP) coupled with massively parallel short-read sequencing (seq) ChIP-seq datasets. PeakRanger provides high performance on extensive data sets by taking advantage of the MapReduce parallel environment. PeakRanger improves recognition of extremely closely-spaced peaks improving spatial accuracy in identifying the exact location of binding events and improving the run time by exploiting the parallel environment provided by a Cloud Computing architecture. PeakRanger is written in C++ and can be deployed on Linux, macOS, and Windows.
- STORMSeq (Scalable Tools for Open-source Read Mapping) [36] is a software pipeline for whole-genome and exome sequence data sets. STORMSeq is implemented as AWS Cloud service. STORMSeq presents an intuitive user interface for dealing with reading mapping and variant calling using genomic data.
- VAT (Variant Annotation Tool) [37] is a software package to annotate variants from multiple individual genomes at the transcript level and obtain descriptive statistics across genes and individuals. VAT visualizes different variants, integrating allele frequencies and genotype data, simplifying comparative analysis between distinct groups of individuals. VAT is implemented in C and PHP and it is available as a command-line tool or as a web application. Moreover, VAT can be run as a virtual machine in the AWS Cloud environment. VAT documentation and user guide are available at http://www.vat.gersteinlab.org (accessed on 21 March 2023).

4. Materials and Methods

This section aims to highlight some challenges, security issues, and impediments limiting the spread of the use of Cloud Computing in healthcare corporations. To identify some of the main relevant obstacles limiting the high adoption of Cloud methodologies in healthcare corporations, we searched the online knowledge database PubMed [38], to figure out from the available scientific literature suitable clues to identify possible advice that could help mitigate the current difficulties in the large use of Cloud Computing in healthcare corporations.

The first step regarded the keywords definition to use for selecting relevant manuscripts. The chosen keywords to implement the selection criteria of the manuscript are: cloud computing, healthcare, security, challenges, applications. Table 2 shows the produced queries obtained by combining the keywords and the selected range of publication years in which to search for manuscripts.

In the second step, we defined the inclusion criteria comprising the following: (i) the manuscripts available on PubMed from the 2009, up to the December 2022 meeting the

selected keywords; (ii) all the types of abstracts, manuscripts, conference abstracts, reviews, and letters are eligible if they contain the chosen keywords in the title and are free full text.

Table 2. The table shows the defined queries to identify relevant manuscripts related to Cloud Computing in healthcare.

QueryID	Query	Publication Years Range
Q_1	cloud computing & healthcare	2009–2022
Q_2	cloud computing & healthcare & security	2009–2022
Q_3	cloud computing & healthcare & challenges	2009–2022
Q_4	cloud computing & healthcare & applications	2009–2022

Table 3 reports the number of identified manuscripts in PubMed that apply to the queries contained in Table 2. The results of the queries were analyzed using an in-house Python script, to parse and extract manuscripts' title keywords, computing for each keyword its frequency (excluding from the frequency terms counting articles, prepositions, adverbs etc). Finally, keyword frequency is used to produce the word cloud diagram shown in Figure 1.

Table 3. The table shows the total number of eligible PubMed manuscripts matching the defined queries.

QueryID	TotManuscripts	TotFreeFullText
Q_1	668	408
Q_2	237	151
Q_3	184	120
Q_4	273	186

Figure 1 presents the results of query Q_1 in the form of word cloud diagram.

Figure 1. Figure shows the query Q_1 results as word cloud diagram.

Figure 2 shows the publication growth trend of manuscripts concerning the use of Cloud Computing in healthcare.

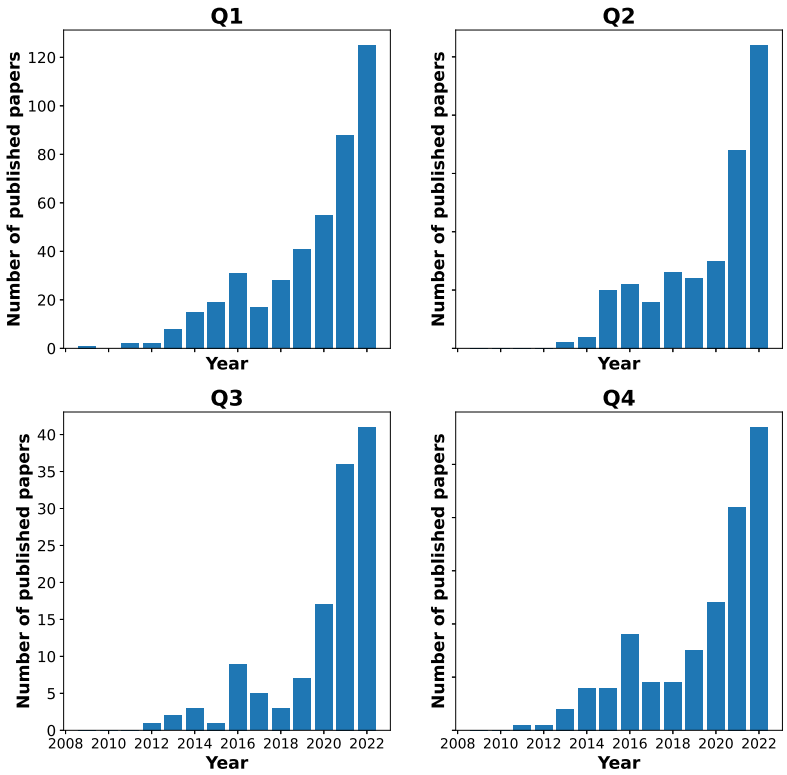

Figure 2. Figure shows the growth trend of Cloud Computing in healthcare starting from 2008 up to December 2022. Q_1 presents the growth per year of manuscripts dealing with cloud computing in healthcare. Q_2 shows the trends per year of the manuscripts focused on security issues in Cloud Computing especially within Cloud Computing in healthcare. Q_3 shows the growth of manuscripts focused on the challenges to be faced in Cloud Computing for healthcare. Finally, Q_4 provides an overview of the growth per year of Cloud application for healthcare.

To highlight the difficulties of adopting Cloud Computing in the healthcare sector, we will analyze the results obtained from the queries represented graphically using piecharts. Figure 3 shows the results of query Q_1.

Q_1 contains the following keywords cloud computing and healthcare, resulting in 67 keywords extracted (for readability reasons, the piechart visualize the first 30 keywords) from the titles of the scientific articles selected using the previously defined criteria concerning the use of Cloud Computing in Healthcare. Analyzing the frequency of keywords identified by query Q_1 shown in Figure 3, it is worth noting that many terms are related to healthcare, which could lead to misleading conclusions concerning the use of Cloud Computing in healthcare, considering that keywords such as security and privacy occupy the 35th and 38th position, respectively.

Query Q_2 adds the keyword security to query Q_1, extracting from scientific works compatible with the selection criteria 17 keywords. Adding the keyword security restricts the selection and search range of the query. In fact, from the result of Q_2 shown in Figure 4, it is possible to notice that the keywords related to security and privacy now leap respectively into 5th, 6th, and 8th position, highlighting the importance of the concepts of security and privacy in the various areas of use of the Cloud and, in particular, in the health sector.

Query Q_3, composed of keywords cloud computing, healthcare and challenges, locates 20 keywords, as shown in Figure 5.

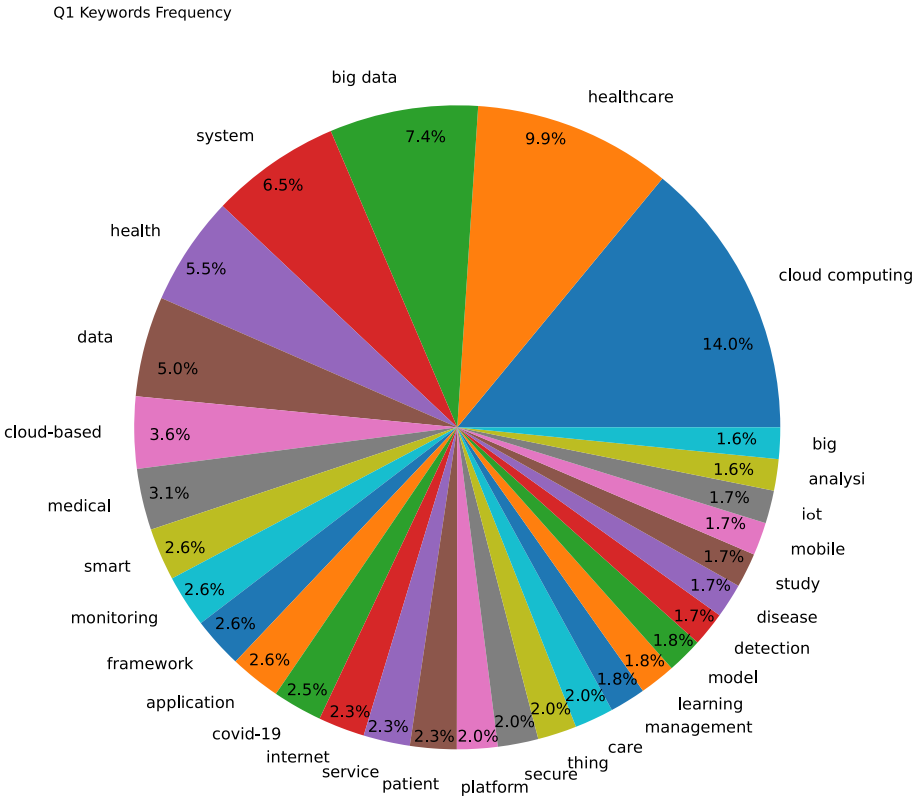

Figure 3. Figure shows the keyword frequency produced from query Q_1. To improve legibility, the percentage values have been truncated to the first value after the decimal point.

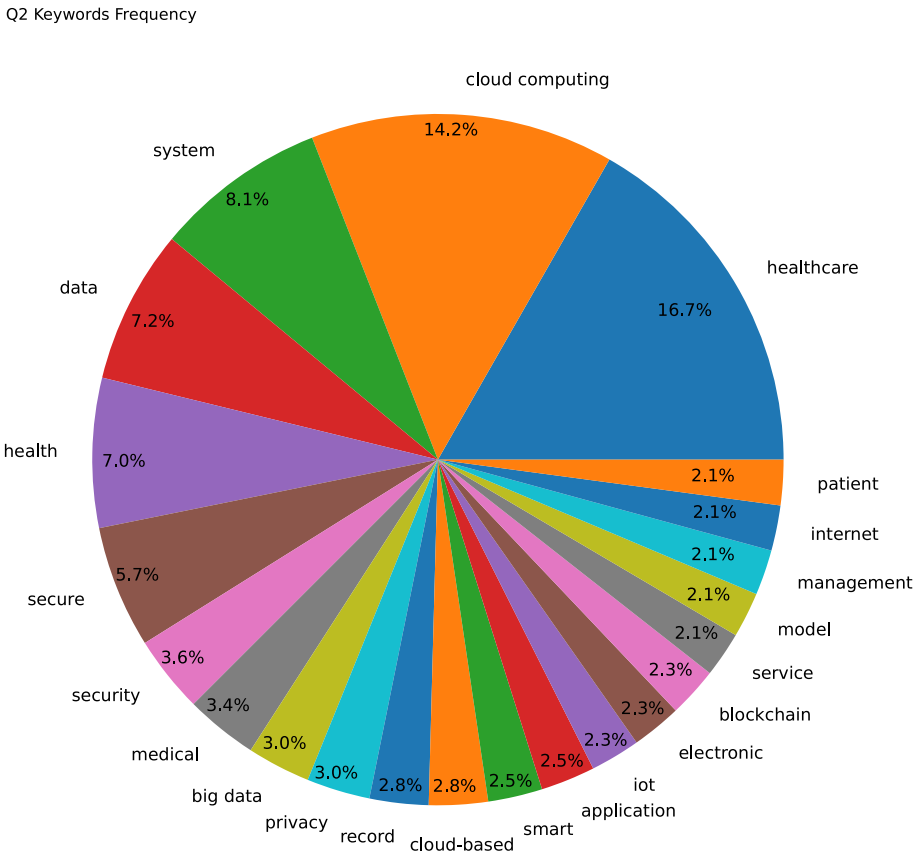

Figure 4. Figure shows the keyword frequency produced from query Q_2. To improve legibility, the percentage values have been truncated to the first value after the decimal point.

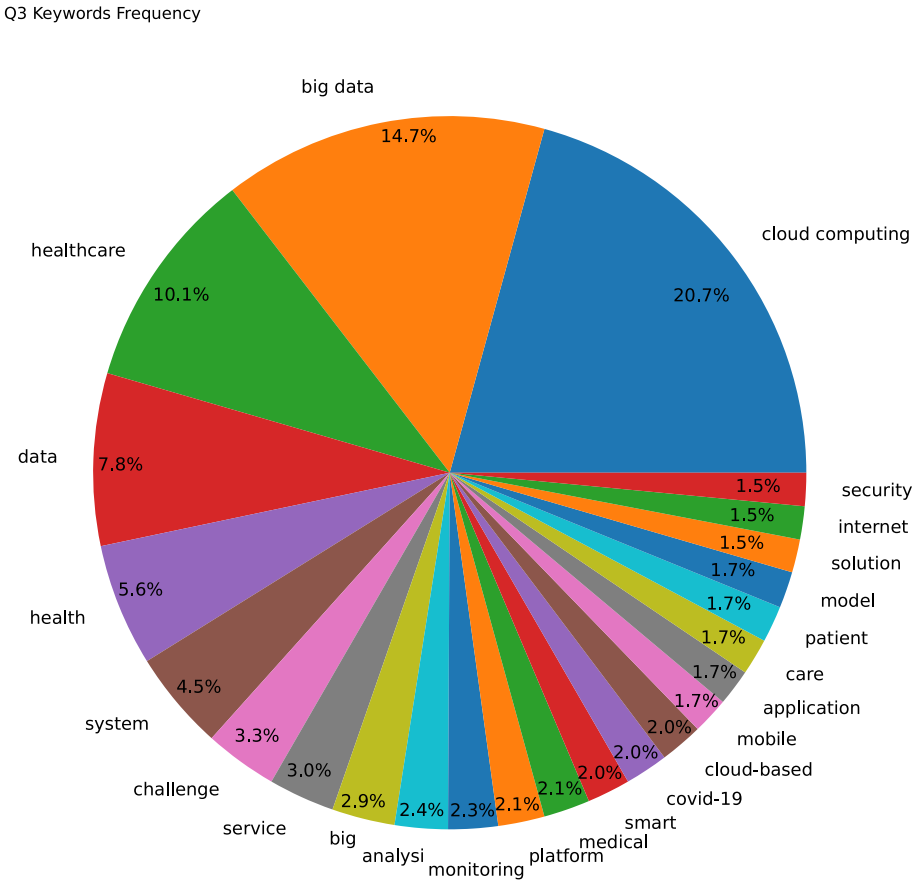

Figure 5. Figure shows the keyword frequency produced from query Q_3. To improve legibility, the percentage values have been truncated to the first value after the decimal point.

In particular, challenges occupies the 5th position, highlighting that the use of the Cloud in the healthcare sector must overcome various challenges, particularly related to the sensitive aspects of the data to be handled. Figure 6 displays the frequency of keywords extracted from query Q_4.

From the analysis of Figure 6 security occupies the 19th position, while privacy does not appear in the list of frequent keywords, introducing biases in the interpretation of the results, suggesting that the existing Cloud Computing applications are mainly aimed at sectors other than healthcare, as less stringent privacy requirements regulate them. In light of these conclusions, decisions regarding the relevant scientific papers to be analyzed were made using the intersection of the results produced by the four queries as a selection criterion.

To limit the manuscripts investigation, we computed the intersection among the results obtained from the four queries performed in PubMed. Figure 7 shows the intersection among the manuscripts' keywords retrieved from each query). The manuscripts intersection was computed using Venny 2.0 [39] a web application used to draw Venn diagrams.

Analysing Figure 7 it is wort noting that the intersection among the four queries contains 27 manuscripts. According to the eligibility criteria, 21 manuscripts have been excluded since they are not explicitly related to Cloud Computing. Finally, only the 6 manuscripts meeting the eligibility criteria have been assessed.

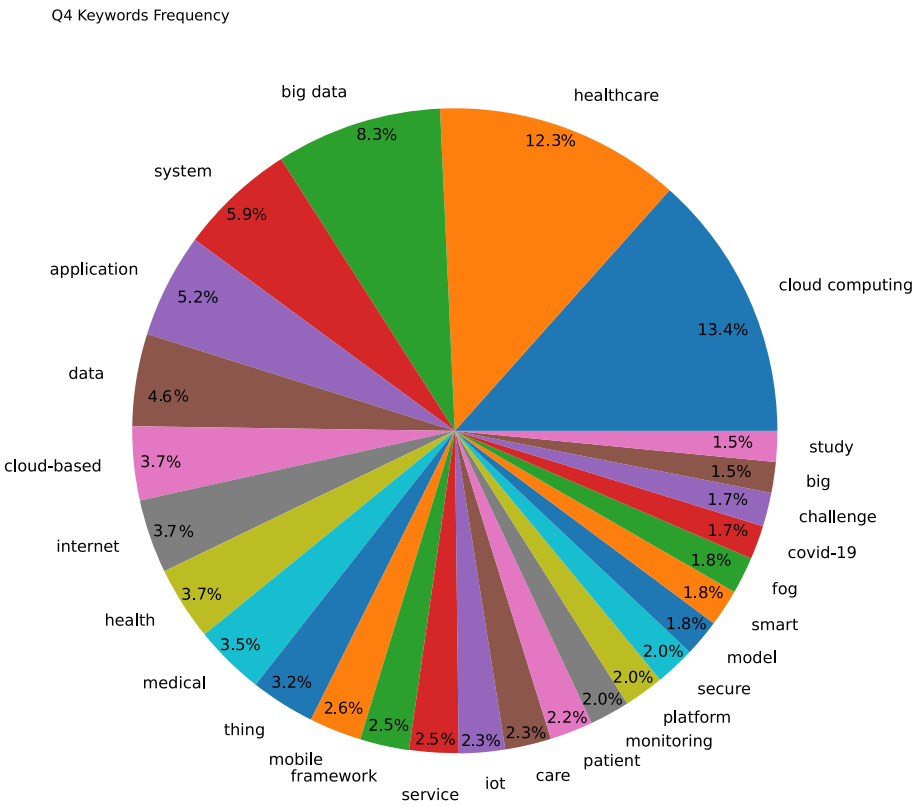

Figure 6. Figure shows the keywords' frequency produced from query Q_4. To improve legibility, the percentage values have been truncated to the first value after the decimal point.

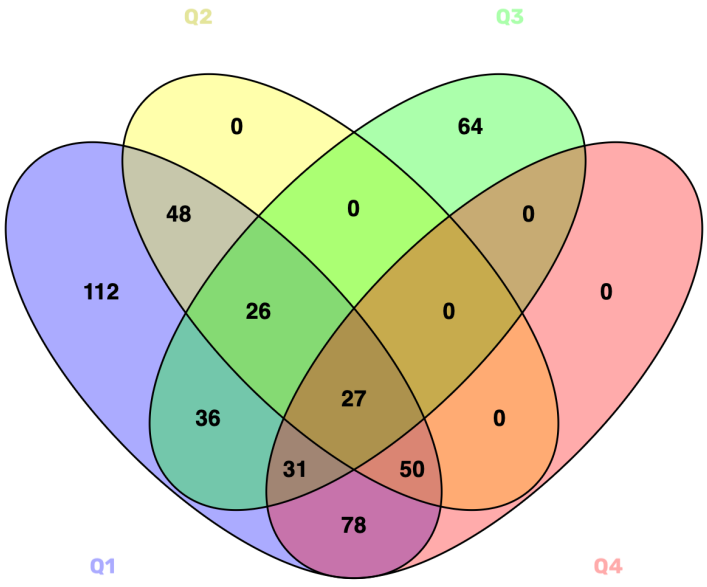

Figure 7. Figure shows the intersection among the manuscripts' keywords retrieved from each query.

5. Discussion

Although Cloud Computing is a consolidated technology in computational and storage resources, with the explicit goals of reducing operating costs and improving results in many scientific domains, Cloud Computing is slowly gathering steam in healthcare despite those premises. This impasse may be due to the critical challenges to face, such as encryption, user identification, storage, access, etc.

Patient clinical information is now collected in Electronic Medical Records (EMRs), even known as Electronic Health Records (EHRs). Using Cloud tools to analyze and share

EMRs data can improve the performance of healthy corporations. Cloud services lowered the cost of care, improved outcomes, and increased customer/patient loyalty and satisfaction while yielding growth and profitability. At the same time, EMRs data must be stored and handled according to well-defined privacy and security rules [40]. Cloud environments must face several challenges in data handling, notably the native heterogeneity of healthcare data and the need to harmonize data sets from different healthcare organizations. Cloud storage is the ideal solution for storing data from different healthcare organizations. It can spur multi center data analysis, data summarization, integration, and harmonization, contributing to new knowledge, improving clinical trials, and developing new drugs. The need for suitable integration and harmonization functions hamper the collaboration between healthcare institutions. Traditional harmonization and integration methods are ineffective with healthcare data. In [41], authors present HarmonicSS, a PaaS Cloud Computing model encouraging collaboration among multiple organizations, providing several data harmonization functions based on semantic data models to identify concepts automatically without a human supervisor. In addition, HarmonicSS provides trustworthy AI models based on the Cloud Federated environment, allowing secure, legal, and ethical uploads compliant with HL7 standards ideal for the healthcare domain. The ubiquity of EMRs in recent years through Cloud Computing could lead to the wide use of artificial intelligence (AI) [42] to analyze these vast amounts of data. AI tools are unhurriedly supplanting humans in many application domains, such as deciding who should get a loan, hiring new workers, and supporting doctors in clinical reporting, decisions, and treatments design. The use of AI in fields where data-driven algorithmic decision-making may affect human life, e.g., healthcare, raises concerns regarding their reliability [43]. Indeed, since AI is a data-driven decision-making tool, using unbalanced, poor, or misleading data sets can increase the probability that these tools could be biased. Improving AI reliability can increase its adoption in healthcare environments. Thus, the challenge is establishing an end-to-end Cloud Computing service able to increase the reliability of AI tools. A potential Cloud Computing service includes the following steps: data acquisition, preprocessing, and AI model training. A possible strategy for increasing end-to-end reliability consists of the following: data labeling, which allows one to figure out the quality of data for the application; results aggregation to simplify the quality assessment; and finally, detection of unbalanced groups, which enables one to obtain more accurate and expressive knowledge models. Hence, the combination of Cloud Computing and reliable AI tools provides Cloud services that can help to increase the adoption of Cloud Computing services in healthcare organizations.

EMR data storage in Cloud repositories throws security problems, such as protecting patients' personal information [44]. Cloud providers can protect EMR-sensitive information by employing noncryptographic techniques such as anonymization and splitting [45]. Data anonymization [46] is a privacy technique to protect a user's personal information, hiding sensitive information that could reveal the identity. Data anonymization can be accomplished by applying various methods, such as removing or hiding identifiers or attributes. The primary intent of data anonymization is to obscure the person's identity in any way. Data splitting divides sensitive data into smaller chunks, distributing those smaller units to distinct storage locations to protect it from unauthorized access. In this manner, data anonymization and splitting protect patients' sensitive information without compromising Cloud Computing performance since data retrieval is accomplished without further computations such as decryption. Noncryptographic techniques provide a basic security level for Cloud environments because intruders can obtain access to complete sensitive information in case of a breach.

Thus, using cryptography [47] can improve Cloud environment security. Cryptography is a fundamental and widely used approach for hiding and securing classified information. Cryptography transforms the raw data into ciphertext using encryption algorithms to protect data during network transfer and storage. Today, cryptography is employed to pursue different targets, such as data confidentiality and integrity. Due to the

increased data violations in the last few years, some Cloud service providers are moving toward cryptographic techniques to attain more safety. In [48], Hassan et al. discuss the relevance of synthesizing, classifying, and identifying different data protection methodologies. Although cryptography increases the security and trust of Cloud environments, it negatively affects Cloud environments' performance. Users want to retrieve their data stored in a Cloud database. Searching for encrypted data is a crucial element of cryptography because every user who stores sensitive data in a local or Cloud database wants to retrieve it. Data retrieving is completed by searching sensitive data through queries. Consequently, the procedure of retrieving data is complicated, since it is not possible to carry out computation on encrypted data without ever decrypting the content.

Cryptography approaches [49,50] are classified into Asymmetric and Symmetric. Asymmetric cryptography [51], also known as a public key, is a technique that uses a couple of keys to encrypt and decrypt information. A key in the pair is public that, as the name implies, can be distributed without affecting security. At the same time, the second key in the pair is private and known exclusively to the owner. In this approach, anyone can use the public key to encrypt messages, but only the paired private key can decrypt those encrypted messages. Public keys are usually stored in digital certificates, which allows them to be easily and securely shared. Private keys are not shared and must be held by users in suitable software systems or hardware, such as USB tokens. Symmetric cryptography [52], also known as a secret key, is a technique that uses a single key for encryption and decryption purposes. In symmetric cryptography, the secret key is private and a secure channel is required to distribute it. This requirement has proved challenging to maintain, representing the main weaknesses of this cryptographic schema. Hence, the key length can mitigate this weakness. In fact, the longer the key, the more secure the communication will be. For instance, to force a key of 128-bit with the computing power of current computers would take millions of years, a sufficient time to guarantee a secure outcome of communications. In asymmetric cryptography, on the other hand, public keys can be distributed on a (possibly) insecure channel, while private keys are generated locally without requiring to be transmitted. This public distribution allows for encrypted and authenticated communications between parties who have not previously met or exchanged information. To summarize, given their different nature, the two types of encryptions are used in purely different fields. Symmetric encryption is used to encrypt files and data when it is necessary to transfer large blocks of information, as well as during data transmission in HTTPS. In contrast, asymmetric cryptography is used in encryption and authentication procedures such as digital signatures. In this regard, healthcare corporations can use symmetric cryptography to achieve more security when sharing data through the network and choose asymmetric cryptography to provide secure authentication procedures to limit access to the stored sensitive information exclusively to the legitimate owner.

Blockchain technology is well known and used in cryptocurrency, safety, and trust management, making it suitable even for Cloud Computing services in healthcare. In [53], Rahmani et al. discussed the issues related to security breaches that occurred in Cloud platforms. Trust handling is critical for delivering secure and trustworthy service to users. The traditional trust-handling protocols in Cloud Computing are centralized, resulting in single-point failure. Hence, Rahmani et al. propose as a solution the use of Blockchain in Cloud domains, e.g., healthcare, that requires trust and trustworthiness in several aspects. An essential feature of Blockchain is the decentralization of the trust model that produces a trust Cloud environment. In [54], Ismail et al. present the limitations of a healthcare system based on either Cloud or Blockchain, highlighting the importance of implementing an integrated Blockchain-Cloud (BcC) system for further improve the Blockchain decentralization and, consequently, the Cloud environment trust.

The Internet of things (IoT) is a paradigm that allows different objects, e.g., intelligent entities and sensors, to communicate with each other on the Internet network. The IoT provides several benefits in many domains, from home to private and public corporations and government institutions. The IoT provides endless opportunities to connect homes,

wearable devices, smart cities, and how patients interact with healthcare corporations. Smart devices, sensors, and wearables, even called smart-objects, are changing how personal care is delivered. Sensors like wearable trackers, e.g., smartwatches and bands, enable automatic self-monitoring and controlling health conditions such as hypertension and blood pressure. Patients can monitor their health status and, if necessary, communicate with their medical doctors to receive expert care directions, improving the quality of their medical care. In [55], the authors provide a picture of how IoT device use changes health care delivery. Thus, despite the above benefits, many issues must be considered, especially data security and privacy, because sensitive patient and hospital information are exchanged over the Internet.

In [56], Kibiwott et al. argue that if the IoT data are far from the owner's physical domain, privacy and security cannot be ensured. In this regard, Kibiwott et al. propose adopting attribute-based signcryption (ABSC) to mitigate security issues and protect sensitive data. ABSC cryptographic properties include fine-grained access control, authentication, confidentiality, and data owner privacy.

To bypass exchanging sensitive information over the network and preventing in this way to face data security and privacy issues, it is possible to use Edge Computing. Edge Computing is a novel programming model aiming to keep the computing step as near to the data source as possible, enabled by the availability of novel devices such as NVIDIA Jetson [57,58]. Moreover, the computation close to the data source guarantees a faster response with low latency, one of the essential requirements in decision-making or mission-critical processes. In [59], the authors present E-ALPHA (Edge-based Assisted Living Platform for Home cAre), which supports both Edge and Cloud Computing paradigms to design innovative Ambient Assisted Living (AAL) services in scenarios of different scales. E-ALPHA flexibly combines Edge and Cloud, assisting users in the preliminary assessment. In particular, it helps to determine the desired performance of the service. Next, it assists users in configuring applications or platforms for real deployment. IoT devices are continuously increasing in many domains, such as scientific, corporate, and domestic, presenting new challenges in the real-time elaboration of these vast amounts of different types of data produced. For these reasons, many initiatives investigating the deployment of architecture-based Edge Computing services and their impact on performance and cost are arising [60]. Moreover, Edge Computing, Machine Learning and Data Mining can put forward the analysis of IoT data based on Edge Computing, Machine Learning, and Deep Learning [61]. In [57], the authors present an approach based on Machine Learning and Edge Computing to diagnose early-stage cancer, allowing efficient and fast analysis without compromising the privacy of sensitive information. In [62] authors proposed EdgeMiningSim, a methodology aimed at IoT domain experts, for creating descriptive or predictive models to take actions in the IoT field.

In [63], Bertuccio et al. describe ReportFlow as an application to transfer sensitive data over the Public Cloud, speeding and simplifying the medical report process of EEGs. ReportFlow exploits the Role-Based Access Control (RBAC) to limit system access only to authorized users. ReportFlow deals with all cryptographic activities, managing certificates and checking their validity using OpenSSL, an open-source general-purpose cryptography library. Public keys and other information are held in specific folders on the Cloud. ReportFlow encrypts the data through a Triple Data Encryption Symmetric Algorithm (Triple DES or 3DES). Finally, Mehrtak et al. in [64] investigated several manuscripts to highlight the importance of accurately determining security challenges and their proper solutions that are fundamental for both Cloud Computing providers and corporations using Cloud services.

To summarize, the slow adoption of Cloud solutions in healthcare organizations could be related to the types of data produced by healthcare organizations. Healthcare data contain sensitive and confidential information about patients, requiring special handling. Thus, it is mandatory to develop special protocols and methods able to protect healthcare data that will be transferred through unsecured channels, i.e., through the internet network, up to the storage, analysis, retrieving etc.

6. Tips to Effectively Use Cloud Computing in Healthcare

This section provides some tips to facilitate the choice of the ideal Cloud Computing provider and how an user can deal with Cloud Computing to meet all the law requirements for healthcare corporations. Customers should choose a Cloud service holding stringent HIPAA and HITECH Act security requirements. Meet HIPAA and HITECH Act security requirements allow to limitate the common vulnerabilities that lead to breaches in security, implementing natively security protocols such as data encryption, multi-factor authentication, intrusion detection, and prevention. In this scenario, Cloud services will be more secure against data breaches, tampering, loss, and damage than on-premise data centers. Consumers would put data in the Cloud storage to create a central point of sharing, intending to promote interoperability. Interoperability can be achieved only if the Cloud provider supplies access to all services to constrained authenticated and authorized entities. Restricted data access with authentication and qualification reduces inappropriate and forbidden data changes. In this manner, data remain intact, secure, and adequately protected. Further consideration should regard the data transfer from local to Cloud repositories. Before uploading sensitive data, users must protect data using cryptography approaches like the HL7 standard and a secure channel like the https. In this way, data remain safe and adequately protected even during the transfer. Before uploading, data summarization, aggregation, and harmonization, in conjunction with encryption, promote secure data analysis, even using advanced AI tools. In this manner, it is possible to prevent AI tools from misusing sensitive information that can harm privacy by introducing biases in the outcomes, contributing to increasing AI reliability. Finally, before choosing a Cloud Computing provider, one must identify the geographic position of Cloud facilities since the security principles depend on the laws of the State and the corresponding legal jurisdiction where it will be held. For programmers, Cloud platforms provide a much faster and more secure method of developing and deploying collaborative, customized, and analytical workflows for dealing with heterogeneous data. In addition, Cloud platforms provide all the software tools, libraries, and APIs to design and develop robust services concerning security threats because the Cloud infrastructure has already been certified. Moreover, the Cloud also reduces the costs associated with maintaining existing infrastructure. In this manner, the internal IT resources can be concentrated on specific tasks rather than handling or maintaining data center hardware.

7. Conclusions

In this paper, we highlighted the importance of identifying Cloud security issues essential to defend patient privacy, complying with healthcare laws and ensuring that only authorized persons can access patients' sensitive data. Thus, the spread use of Cloud in healthcare could be enhanced by providing trusted Cloud architectures and services where the privacy and security of all data types is explicitly ensured, rendering information misuse impossible.

Author Contributions: Conceptualization, G.A.; methodology, G.A.; investigation, G.A.; writing, review and editing, G.A. and M.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially funded by the Data Analytics Research Center, and the "Cultura Romana del Diritto e Sistemi Giuridici Contemporanei" Research Center, Catanzaro, Italy.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MC	Molecular Biology
HT	High-throughput
AWS	Amazon Web Services
BPaaS	Business Process as a Service
CaaS	Connectivity as a Service
ChIP	Chromatin immunoprecipitation
ChiPseq	Short read sequencing
DaaS	Data as a Service
DNA	DeoxyriboNucleic Acid
EMR	Elastic MapReduce
EMR	Hectronic medical record
GPU	raphics processing units
HIPAA	Health Insurance Portability and Accountability Act
HPC	High-Performance Computing
IaaS	Infrastructure as a Service
IDaaS	Identity as a Service
IT	Information Technology
MPI	Message-Passing Interface
NGS	Next-Generation Sequence
PaaS	Platform as a Service
RMAP	short read-mapping program
RNA-seq	RNA sequence
SaaS	Software as a Service
scRNA-seq	Single-cell RNA-sequence
SNP	Single Nucleotide Polymorphism
STORMSeq	Scalable Tools for Open-source Read Mapping
VAT	Variant Annotation Tool
VM	Virtual machines

References

1. Ahn, A.C.; Tewari, M.; Poon, C.S.; Phillips, R.S. The limits of reductionism in medicine: Could systems biology offer an alternative? *PLoS Med.* **2006**, *3*, e208. [CrossRef] [PubMed]
2. Loscalzo, J.; Barabasi, A.L. Systems biology and the future of medicine. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2011**, *3*, 619–627. [CrossRef] [PubMed]
3. Vailati-Riboni, M.; Palombo, V.; Loor, J.J. What are omics sciences? In *Periparturient Diseases of Dairy Cows*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 1–7.
4. Mardis, E.R. Next-generation DNA sequencing methods. *Annu. Rev. Genom. Hum. Genet.* **2008**, *9*, 387–402. [CrossRef] [PubMed]
5. Shendure, J.; Balasubramanian, S.; Church, G.M.; Gilbert, W.; Rogers, J.; Schloss, J.A.; Waterston, R.H. DNA sequencing at 40: Past, present and future. *Nature* **2017**, *550*, 345–353. [CrossRef]
6. D'Adamo, G.L.; Widdop, J.T.; Giles, E.M. The future is now? Clinical and translational aspects of "Omics" technologies. *Immunol. Cell Biol.* **2021**, *99*, 168–176. [CrossRef]
7. Schneider, M.V.; Orchard, S. Omics technologies, data and bioinformatics principles. *Bioinform. Omics Data* **2011**, *719*, 3–30.
8. Clarke, L.; Glendinning, I.; Hempel, R. The MPI message passing interface standard. In *Programming Environments for Massively Parallel Distributed Systems*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 213–218.
9. Kim, W. Cloud computing: Today and tomorrow. *J. Object Technol.* **2009**, *8*, 65–72. [CrossRef]
10. Dillon, T.; Wu, C.; Chang, E. Cloud computing: Issues and challenges. In Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, WA, Australia, 20–23 April 2010; pp. 27–33.
11. Pautasso, C.; Wilde, E. RESTful web services: Principles, patterns, emerging technologies. In Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC, USA, 26–30 April 2010; pp. 1359–1360.
12. Cusumano, M. Cloud computing and SaaS as new computing platforms. *Commun. ACM* **2010**, *53*, 27–29. [CrossRef]
13. Pahl, C. Containerization and the paas cloud. *IEEE Cloud Comput.* **2015**, *2*, 24–31. [CrossRef]
14. Bhardwaj, S.; Jain, L.; Jain, S. Cloud computing: A study of infrastructure as a service (IAAS). *Int. J. Eng. Inf. Technol.* **2010**, *2*, 60–63.
15. Woitsch, R.; Utz, W. Business process as a service (BPaaS). In Proceedings of the Conference on e-Business, e-Services and e-Society, Delft, The Netherlands, 13–15 October 2015; pp. 435–440.

16. Rajesh, S.; Swapna, S.; Reddy, P.S. Data as a service (daas) in cloud computing. *Glob. J. Comput. Sci. Technol.* **2012**, *12*, 25–29.
17. Ni, Y.; Xing, C.L.; Zhang, K. Connectivity as a service: Outsourcing Enterprise connectivity over cloud computing environment. In Proceedings of the 2011 International Conference on Computer and Management (CAMAN), Wuhan, China, 19–21 May 2011; pp. 1–7.
18. Ducatel, G. Identity as a service: A cloud based common capability. In Proceedings of the 2015 IEEE Conference on Communications and Network Security (CNS), Florence, Italy, 28–30 September 2015; pp. 675–679.
19. Krampis, K.; Booth, T.; Chapman, B.; Tiwari, B.; Bicak, M.; Field, D.; Nelson, K.E. Cloud BioLinux: Pre-configured and on-demand bioinformatics computing for the genomics community. *BMC Bioinform.* **2012**, *13*, 42. [CrossRef]
20. Agapito, G.; Cannataro, M.; Guzzi, P.H.; Marozzo, F.; Talia, D.; Trunfio, P. Cloud4SNP: Distributed analysis of SNP microarray data on the cloud. In Proceedings of the International Conference on Bioinformatics, Computational Biology and Biomedical Informatics, Washington, DC, USA, 22–25 September 2013; pp. 468–475.
21. Guzzi, P.H.; Agapito, G.; Di Martino, M.T.; Arbitrio, M.; Tassone, P.; Tagliaferri, P.; Cannataro, M. DMET-analyzer: Automatic analysis of Affymetrix DMET data. *BMC Bioinform.* **2012**, *13*, 258. [CrossRef]
22. Marozzo, F.; Talia, D.; Trunfio, P. A Cloud Framework for Big Data Analytics Workflows on Azure. In *Proceedings of the Post-Proceedings of the High Performance Computing Workshop 2012*; Catlett, C., Gentzsch, W., Grandinetti, L., Joubert, G., Vazquez-Poletti, J.L., Eds.; IOS Press: Cetraro, Italy, 2013; Volume 23, pp. 182–191, ISBN 978-1-61499-321-6.
23. Marozzo, F.; Talia, D.; Trunfio, P. Using clouds for scalable knowledge discovery applications. In Proceedings of the European Conference on Parallel Processing, Aachen, Germany, 26–30 August 2013; pp. 220–227.
24. Schatz, M.C. CloudBurst: Highly sensitive read mapping with MapReduce. *Bioinformatics* **2009**, *25*, 1363–1369. [CrossRef]
25. Dean, J.; Ghemawat, S. MapReduce: Simplified data processing on large clusters. *Commun. ACM* **2008**, *51*, 107–113. [CrossRef]
26. Afgan, E.; Chapman, B.; Taylor, J. CloudMan as a platform for tool, data, and analysis distribution. *BMC Bioinform.* **2012**, *13*, 315. [CrossRef]
27. Afgan, E.; Lonie, A.; Taylor, J.; Goonasekera, N. CloudLaunch: Discover and deploy cloud applications. *Future Gener. Comput. Syst.* **2019**, *94*, 802–810. [CrossRef]
28. Afgan, E.; Baker, D.; Coraor, N.; Chapman, B.; Nekrutenko, A.; Taylor, J. Galaxy CloudMan: Delivering cloud compute clusters. In Proceedings of the BMC Bioinformatics, Boston, MA, USA, 9–10 July 2010; Volume 11, pp. 1–6.
29. Langmead, B.; Schatz, M.; Lin, J.; Pop, M.; Salzberg, S. Searching for snps with cloud computing. *Genome Biol.* **2009**, *10*, R134. [CrossRef]
30. Li, R.; Li, Y.; Fang, X.; Yang, H.; Wang, J.; Kristiansen, K.; Wang, J. SNP detection for massively parallel whole-genome resequencing. *Genome Res.* **2009**, *19*, 1124–1132. [CrossRef]
31. Jourdren, L.; Bernard, M.; Dillies, M.A.; Le Crom, S. Eoulsan: A cloud computing-based framework facilitating high throughput sequencing analyses. *Bioinformatics* **2012**, *28*, 1542–1543. [CrossRef]
32. Lehmann, N.; Perrin, S.; Wallon, C.; Bauquet, X.; Deshaies, V.; Firmo, C.; Du, R.; Berthelier, C.; Hernandez, C.; Michaud, C.; et al. Eoulsan 2: An efficient workflow manager for reproducible bulk, long-read and single-cell transcriptomics analyses. *bioRxiv* **2021**. [CrossRef]
33. Ehwerhemuepha, L.; Gasperino, G.; Bischoff, N.; Taraman, S.; Chang, A.; Feaster, W. HealtheDataLab—A cloud computing solution for data science and advanced analytics in healthcare with application to predicting multi-center pediatric readmissions. *BMC Med. Informatics Decis. Mak.* **2020**, *20*, 115. [CrossRef] [PubMed]
34. Liu, L.; Chen, W.; Nie, M.; Zhang, F.; Wang, Y.; He, A.; Wang, X.; Yan, G. iMAGE cloud: Medical image processing as a service for regional healthcare in a hybrid cloud environment. *Environ. Health Prev. Med.* **2016**, *21*, 563–571. [CrossRef] [PubMed]
35. Feng, X.; Grossman, R.; Stein, L. PeakRanger: A cloud-enabled peak caller for ChIP-seq data. *BMC Bioinform.* **2011**, *12*, 139. [CrossRef]
36. Karczewski, K.J.; Fernald, G.H.; Martin, A.R.; Snyder, M.; Tatonetti, N.P.; Dudley, J.T. STORMSeq: An open-source, user-friendly pipeline for processing personal genomics data in the cloud. *PloS ONE* **2014**, *9*, e84860. [CrossRef]
37. Habegger, L.; Balasubramanian, S.; Chen, D.Z.; Khurana, E.; Sboner, A.; Harmanci, A.; Rozowsky, J.; Clarke, D.; Snyder, M.; Gerstein, M. VAT: A computational framework to functionally annotate variants in personal genomes within a cloud-computing environment. *Bioinformatics* **2012**, *28*, 2267–2269. [CrossRef]
38. Roberts, R.J. PubMed Central: The GenBank of the published literature. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 381–382. [CrossRef]
39. Oliveros, J.C. VENNY. An Interactive Tool for Comparing Lists with Venn Diagrams. 2007. Available online: http://bioinfogp.cnb.csic.es/tools/venny/index.html (accessed on 15 March 2023).
40. Calabrese, B.; Cannataro, M. Cloud computing in healthcare and biomedicine. *Scalable Comput. Pract. Exp.* **2015**, *16*, 1–18. [CrossRef]
41. Pezoulas, V.C.; Goules, A.; Kalatzis, F.; Chatzis, L.; Kourou, K.D.; Venetsanopoulou, A.; Exarchos, T.P.; Gandolfo, S.; Votis, K.; Zampeli, E.; et al. Addressing the clinical unmet needs in primary Sjögren's Syndrome through the sharing, harmonization and federated analysis of 21 European cohorts. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 471–484. [CrossRef]
42. Bukowski, M.; Farkas, R.; Beyan, O.; Moll, L.; Hahn, H.; Kiessling, F.; Schmitz-Rode, T. Implementation of eHealth and AI integrated diagnostics with multidisciplinary digitized data: Are we ready from an international perspective? *Eur. Radiol.* **2020**, *30*, 5510–5524. [CrossRef]
43. Shneiderman, B. Human-centered artificial intelligence: Reliable, safe & trustworthy. *Int. J. Hum. Comput. Interact.* **2020**, *36*, 495–504.

44. Wu, Z.; Xuan, S.; Xie, J.; Lin, C.; Lu, C. How to ensure the confidentiality of electronic medical records on the cloud: A technical perspective. *Comput. Biol. Med.* **2022**, *147*, 105726. [CrossRef]
45. Gkoulalas-Divanis, A.; Loukides, G. *Anonymization of Electronic Medical Records to Support Clinical Analysis*; Springer: Berlin/Heidelberg, Germany, 2012.
46. Majeed, A.; Lee, S. Anonymization techniques for privacy preserving data publishing: A comprehensive survey. *IEEE Access* **2020**, *9*, 8512–8545. [CrossRef]
47. Ayoub, F.; Singh, K. Cryptographic techniques and network security. In *Proceedings of the IEE Proceedings F-Communications, Radar and Signal Processing*; IEEE: Piscataway, NJ, USA, 1984; Volume 7, pp. 684–694.
48. Hassan, J.; Shehzad, D.; Habib, U.; Aftab, M.U.; Ahmad, M.; Kuleev, R.; Mazzara, M. The Rise of Cloud Computing: Data Protection, Privacy, and Open Research Challenges—A Systematic Literature Review (SLR). *Comput. Intell. Neurosci.* **2022**, *2022*, 8303504. [CrossRef]
49. Forouzan, B.A.; Mukhopadhyay, D. *Cryptography and Network Security*; Mc Graw Hill Education Private Limited: New York, NY, USA, 2015; Volume 12.
50. Abood, O.G.; Guirguis, S.K. A survey on cryptography algorithms. *Int. J. Sci. Res. Publ.* **2018**, *8*, 495–516. [CrossRef]
51. Gordon, A.D.; Jeffrey, A. Types and effects for asymmetric cryptographic protocols. *J. Comput. Secur.* **2004**, *12*, 435–483. [CrossRef]
52. Biryukov, A.; Perrin, L. State of the art in lightweight symmetric cryptography. *Cryptol. ePrint Arch.* **2017**. Available online: https://eprint.iacr.org/2017/511 (accessed on 15 March 2023)
53. Rahmani, M.K.I.; Shuaib, M.; Alam, S.; Siddiqui, S.T.; Ahmad, S.; Bhatia, S.; Mashat, A. Blockchain-Based Trust Management Framework for Cloud Computing-Based Internet of Medical Things (IoMT): A Systematic Review. *Comput. Intell. Neurosci.* **2022**, *2022*, 9766844. [CrossRef]
54. Ismail, L.; Materwala, H.; Hennebelle, A. A scoping review of integrated blockchain-cloud (BcC) architecture for healthcare: Applications, challenges and solutions. *Sensors* **2021**, *21*, 3753. [CrossRef]
55. Metcalf, D.; Milliard, S.T.; Gomez, M.; Schwartz, M. Wearables and the Internet of Things for Health: Wearable, Interconnected Devices Promise More Efficient and Comprehensive Health Care. *IEEE Pulse* **2016**, *7*, 35–39. [CrossRef]
56. Kibiwott, K.P.; Zhao, Y.; Kogo, J.; Zhang, F. Verifiable fully outsourced attribute-based signcryption system for IoT eHealth big data in cloud computing. *Math. Biosci. Eng.* **2019**, *16*, 3561–3594. [CrossRef]
57. Barillaro, L.; Agapito, G.; Cannataro, M. Edge-based Deep Learning in Medicine: Classification of ECG signals. In Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Las Vegas, NV, USA, 6–8 December 2022; pp. 2169–2174.
58. Crespo-Cepeda, R.; Agapito, G.; Vazquez-Poletti, J.L.; Cannataro, M. Challenges and Opportunities of Amazon Serverless Lambda Services in Bioinformatics. In Proceedings of the 10th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, Niagara Falls, NY, USA, 7–10 September 2019; pp. 663–668. [CrossRef]
59. Aloi, G.; Fortino, G.; Gravina, R.; Pace, P.; Savaglio, C. Simulation-driven platform for Edge-based AAL systems. *IEEE J. Sel. Areas Commun.* **2020**, *39*, 446–462. [CrossRef]
60. Casadei, R.; Fortino, G.; Pianini, D.; Placuzzi, A.; Savaglio, C.; Viroli, M. A methodology and simulation-based toolchain for estimating deployment performance of smart collective services at the edge. *IEEE Internet Things J.* **2022**, *9*, 20136–20148. [CrossRef]
61. Barillaro, L.; Agapito, G.; Cannataro, M. Scalable Deep Learning for Healthcare: Methods and Applications. In Proceedings of the 13th ACM International Conference on Bioinformatics, Computational Biology and Health Informatics, Northbrook, IL, USA, 7–10 August 2022. [CrossRef]
62. Savaglio, C.; Fortino, G. A simulation-driven methodology for IoT data mining based on edge computing. *ACM Trans. Internet Technol. (TOIT)* **2021**, *21*, 1–22. [CrossRef]
63. Bertuccio, S.; Tardiolo, G.; Giambò, F.M.; Giuffrè, G.; Muratore, R.; Settimo, C.; Raffa, A.; Rigano, S.; Bramanti, A.; Muscarà, N.; et al. ReportFlow: An application for EEG visualization and reporting using cloud platform. *BMC Med. Inform. Decis. Mak.* **2021**, *21*, 7. [CrossRef] [PubMed]
64. Mehrtak, M.; SeyedAlinaghi, S.; MohsseniPour, M.; Noori, T.; Karimi, A.; Shamsabadi, A.; Heydari, M.; Barzegary, A.; Mirzapour, P.; Soleymanzadeh, M.; et al. Security challenges and solutions using healthcare cloud computing. *J. Med. Life* **2021**, *14*, 448. [CrossRef]