



Review

A Comprehensive Survey on Data Utility and Privacy: Taking Indian Healthcare System as a Potential Case Study

Prathamesh Churi ^{1,2}, Ambika Pawar ¹ and Antonio-José Moreno-Guerrero ^{3,*}

¹ Symbiosis Institute of Technology, Symbiosis International Deemed University, Pune 412115, India; prathamesh.churi@ieee.org (P.C.); ambikap@sitpune.edu.in (A.P.)

² School of Technology Management and Engineering, NMIMS University, Mumbai 400056, India

³ Department of Didactics and School Organization, University of Granada, 18011 Ceuta, Spain

* Correspondence: ajmoreno@ugr.es

Abstract: Background: According to the renowned and Oscar award-winning American actor and film director Marlon Brando, “privacy is not something that I am merely entitled to, it is an absolute prerequisite.” Privacy threats and data breaches occur daily, and countries are mitigating the consequences caused by privacy and data breaches. The Indian healthcare industry is one of the largest and rapidly developing industry. Overall, healthcare management is changing from disease-centric into patient-centric systems. Healthcare data analysis also plays a crucial role in healthcare management, and the privacy of patient records must receive equal attention. **Purpose:** This paper mainly presents the utility and privacy factors of the Indian healthcare data and discusses the utility aspect and privacy problems concerning Indian healthcare systems. It defines policies that reform Indian healthcare systems. The case study of the NITI Aayog report is presented to explain how reformation occurs in Indian healthcare systems. **Findings:** It is found that there have been numerous research studies conducted on Indian healthcare data across all dimensions; however, privacy problems in healthcare, specifically in India, are caused by prevalent complacency, culture, politics, budget limitations, large population, and existing infrastructures. This paper reviews the Indian healthcare system and the applications that drive it. Additionally, the paper also maps that how privacy issues are happening in every healthcare sector in India. **Originality/Value:** To understand these factors and gain insights, understanding Indian healthcare systems first is crucial. To the best of our knowledge, we found no recent papers that thoroughly reviewed the Indian healthcare system and its privacy issues. The paper is original in terms of its overview of the healthcare system and privacy issues. **Social Implications:** Privacy has been the most ignored part of the Indian healthcare system. With India being a country with a population of 130 billion, much healthcare data are generated every day. The chances of data breaches and other privacy violations on such sensitive data cannot be avoided as they cause severe concerns for individuals. This paper segregates the healthcare system’s advances and lists the privacy that needs to be addressed first.



Citation: Churi, P.; Pawar, A.; Moreno-Guerrero, A.-J. A Comprehensive Survey on Data Utility and Privacy: Taking Indian Healthcare System as a Potential Case Study. *Inventions* **2021**, *6*, 45. <https://doi.org/10.3390/inventions6030045>

Academic Editors: Majeed Abdul Chien-Hung Liu

Received: 13 May 2021
Accepted: 17 June 2021
Published: 23 June 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Keywords: healthcare; utility; privacy; Indian healthcare; hospitals; EHR; KPI; PDP bill; data breach; attacks



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The healthcare industry is an emerging industry. In recent years, the healthcare industry in developing countries has grown rapidly. In the last few decades, considerable efforts have been taken to integrate information and communication technologies (ICT) into healthcare practices [1,2]. In E-healthcare, the latest technologies are integrated with medical infrastructures, including continuous monitoring and transfer of health-related problems from the patient-centric environment to respective services providers [2–4]. The volume of the data generated in the healthcare industry is rapidly escalating. In healthcare, for highly accurate prediction and early diagnosis of diseases, we must increase healthcare

data utility with the help of various technologies such as machine learning, artificial intelligence (AI), and data analysis.

According to IBM global business services' executive report in 2012, the entire health-care system is being moved from a disease-centric to a patient-centric environment [4,5]. Disease-centric healthcare systems have the following features (Figure 1):

1. The health data are centrally stored according to diseases (type, symptoms, and remedial medicines).
2. Electronic health records (EHRs) are assessments and analyses conducted according to diseases. For example, we can analyze the patient data of the past ten years in a hospital with diabetes, malaria, or another joint disease.
3. Disease-centric databases present less scope for data analysis because they do not focus on individual traits and symptoms. Some diseases are related to a person's behavior, lifestyle, and geographical location. For example, suppose a person is treated for two or three similar diseases from different hospitals in the past ten years. In that case, we must analyze his treatment records, family records, and habits, which are unavailable in the disease-centric database or may be available in heterogeneous databases, wherein a combined analysis is complex. Some data values are spread across multiple datasets maintained separately by hospitals or may have incomplete values, resulting in inappropriate/wrong prediction. The data quality is questionable to be used for analysis.

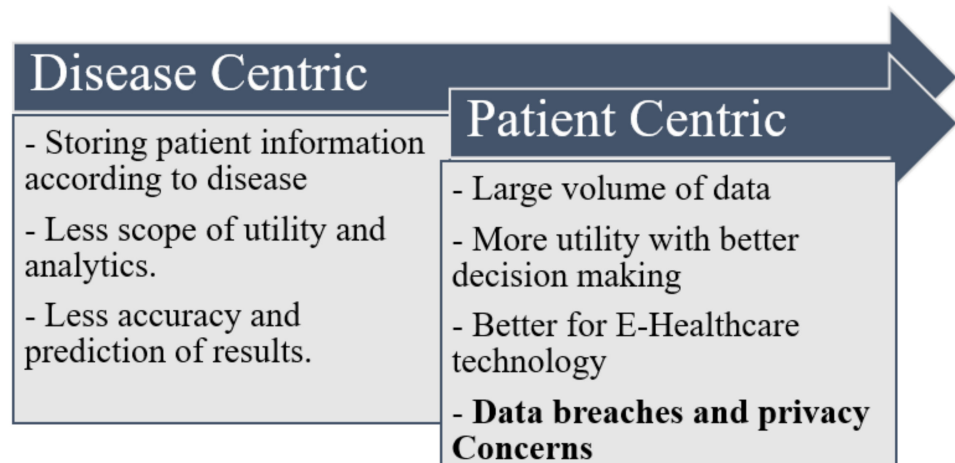


Figure 1. Disease- and patient-centric healthcare systems.

By contrast, patient-centric databases have the following features:

1. The data volume generated is considerably larger and stored according to the individual patient; hence, the data quality and utility are highly satisfactory in appropriate decision making.
2. In E-healthcare, the latest technologies are integrated with medical infrastructures, including continuous monitoring and transfer of health-related problems from the patient-centric environment to respective service providers [2,3,6–9].
3. Because an individual's data are collected from multiple devices and sensors or through sources, maintaining the individual's privacy is challenging.

Security and privacy concerns regarding any type of data are significant problems in the current technology-driven world. With substantial healthcare data for analyses and studies, maintaining privacy is another field that requires further improvement. The data are kept anonymous from users or servers to prevent its misuse. Such health data are processed and stored dynamically at different dynamic locations with various transparencies in the distributed environment. In such a scenario, maintaining data privacy is crucial. Some privacy techniques, namely anonymization, generalization, perturbation,

role-based access control, and encryption, are used to hide data. According to [4], data undergo different phases during its lifecycle: Data storage, transition, transfer, and processing. Existing privacy-preserving techniques remain in the developing stages, and strong privacy protection is still an open study topic.

With the advent of the technologies mentioned above presenting the problems of maintaining data privacy, central questions that remain unaddressed in the healthcare industry field are as follows [10]:

1. Can one pursue high data utility while maintaining acceptable privacy?
2. Because privacy concerns are different for different healthcare organizations, how is the trade-off between privacy protection and data utility balanced for computing?

Figure 2 illustrates the trade-off between data utility and privacy. In the past years, the focus was on maintaining patient privacy and maximizing utility by considering patient privacy [11–13].

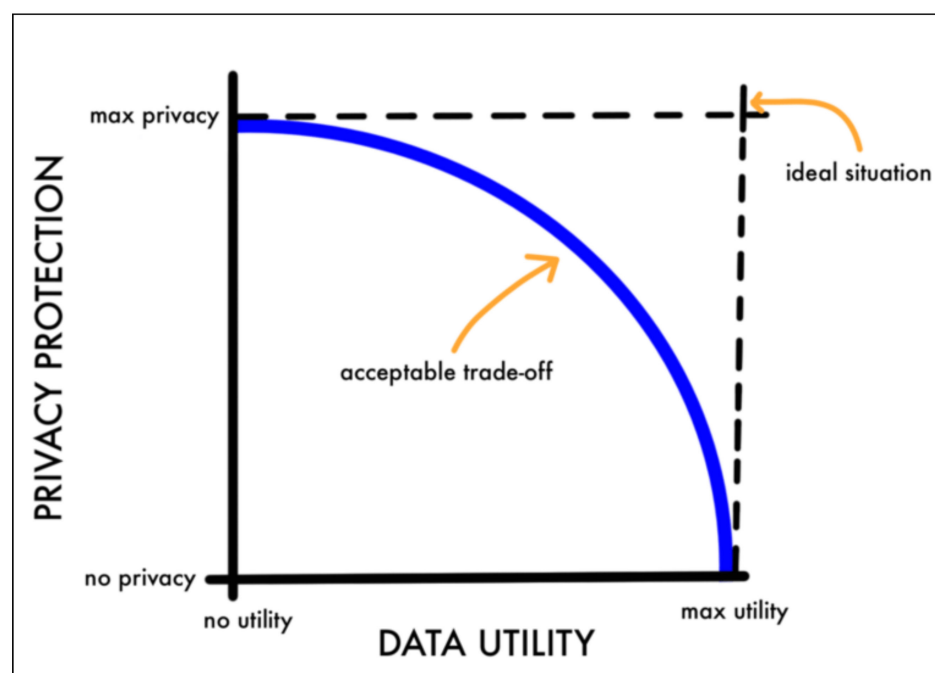


Figure 2. Trade-off between privacy and utility.

With the advent of technology, the number of healthcare markets and assets in India has been increasing every year. India will have a potential healthcare market shortly. Many medical institutes are emerging because of a change in government policies. The Indian government is motivating and encouraging medical colleges to be equipped. Because the Indian healthcare structure is complex and interdependent, technology implementation and addressing privacy problems has always been a big question. Therefore, the contributions of this paper are as follows:

- Provide insights into Indian healthcare systems with applications, trends, and advantages.
- Describe policies that drive Indian healthcare systems
- Specify technological inventions used in Indian healthcare systems.
- List the various privacy issues concerning the Indian healthcare system that needs to be addressed first.

Structure of Paper

The paper mainly presents the utility and privacy of the healthcare data and discusses the utility aspect and privacy problems of Indian healthcare systems (Figure 3). To understand these factors and gain insights, understanding Indian healthcare systems first is

crucial. Section 2 presents overall Indian healthcare systems and world health organization (WHO) indicators that classify sound healthcare systems. Section 3 defines policies that reform Indian healthcare systems. The case study of the NITI Aayog report is presented to explain how reformation occurs in Indian healthcare systems. Section 4 describes healthcare applications in India, wherein the advantages of the healthcare system, trends in healthcare systems, and healthcare startups originated in India are explained. Section 5 presents technologies available for healthcare analytics and Indian papers based on new machine learning and AI strategies. Section 6 addresses the various privacy problems discussed in the literature. Section 7 concludes the study.

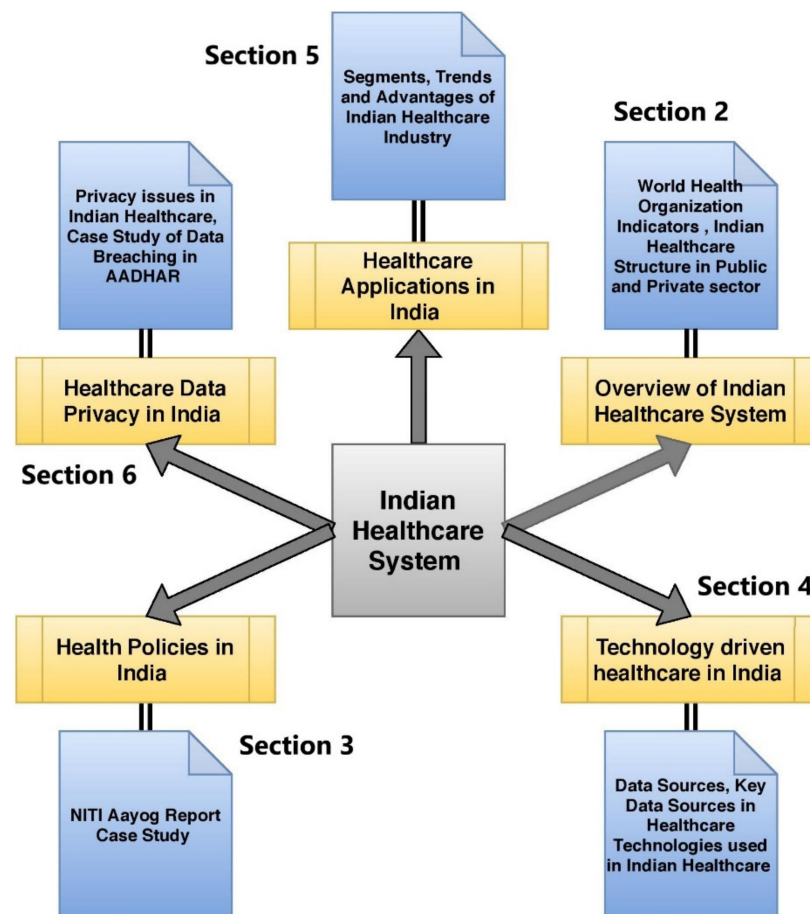


Figure 3. Structure and contribution of the paper.

2. Indian Healthcare Systems: Overview

Indian healthcare systems are divided into two sectors: Public and private. The public sector healthcare is handled by the government and opens for all people. This sector includes super-specialty hospitals equipped with medicines and instruments, which are majorly located in tier I and tier II cities. Additionally, district- and taluka-level hospitals provide healthcare services to the people [11–13]. Primary healthcare centers and village hospitals with low costs are available, which provide affordable services to the people.

The private sector has a similar structure and is generally used by the upper-middle class and upper-class population. The overall cost of healthcare services included in the private sector is higher than that in the public sector. Technological interventions are also more diverse in the private sector than in the public sector. Figure 4 illustrates the detailed structure of the Indian healthcare system. Table 1 presents the difference between the public and private sectors [14]. The differentiation factor is adopted from the WHO

health system themes [14]. WHO presents some descriptive indicators to define a suitable healthcare system.

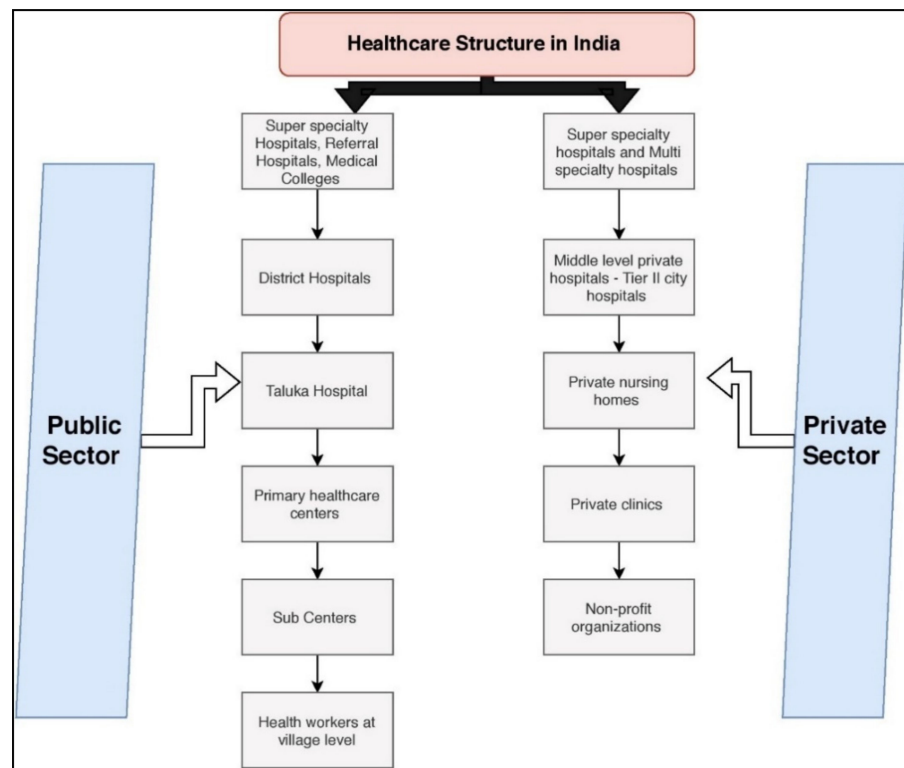


Figure 4. Indian healthcare system.

Table 1. WHO indicators for suitable healthcare systems and Indian context.

Category	Sub-Category	Description and Indicators	The Public Sector in India	The Private Sector in India
Access and response	Availability	24 × seven healthcare service availability to people without any hesitation	Moderate	Good
	Timeliness of service	Less waiting time to initial screening and subsequent testing, providing results, and follow-up	Moderate	Excellent
	Hospitality	Highly responsive feedback system, facility, and maintenance of healthcare system	Moderate	Good
Quality	The comprehensiveness of healthcare services	Availability of all the components of WHO service packages	Poor	Good
	Diagnostic	Accurate diagnosis of retrospective review	Moderate	Excellent
	Management standards	Rate of conformity to international disease-specific management standards	Poor	Good
	Client retention	Rate of failure to follow-up or rate of appropriate patient return	Moderate	Moderate
Outcomes	Treatment success rates	Rate of therapy success, controlling of population characteristics, and delayed presentation	Moderate	Good
	Population coverage	The proportion of the catchment population reached through dedicated campaigns (e.g., vaccination rates)	Excellent	Moderate
	Morbidity	Rate of disability to patients and controlling of population characteristics	Moderate	Less
	Mortality	Rate of patient death and controlling of population characteristics	Moderate	Less

Table 1. Cont.

Category	Sub-Category	Description and Indicators	The Public Sector in India	The Private Sector in India
Accountability, transparency, and regulation.	Data accessibility and quality	Availability of data and appropriate use of indicators and statistics	Poor	Good
	Public health functions	Contribution of healthcare systems to core public health system functions (e.g., reporting of critical diseases and preventative care)	Good	Good
	Reform capacity	Results of quality improvement initiatives	Poor	Good
Fairness and equity	Financial barriers to care	User fees, bribes, and pharmaceutical costs	Very less	High
	Distributive justice	Healthcare availability commensurate with requirements	Moderate	Good
Efficiency	Cost	Absolute dollars spent for a given indication	Very less	High
	Redundancy	Repetition of diagnostic time, testing, supply chains, and therapy delivery	Moderate	Good
	Fragmentation	Separation of core healthcare system functions and generating sluggish management	Poor	Poor
	Delays	The time between the ordering of tests or therapies and their execution	Poor	Good

3. Healthcare in India: Reformation in Policies

Case Study: Healthcare Sector in India: NITI Aayog Report

The Indian government is highly proactive in the healthcare sector and encourages outside investors to invest in the Indian healthcare industry. According to the NITI Aayog statistics [1], the Indian government will increase the public expenditure of Healthcare from 1.1% to 2.5% GDP in the next four years. This finding shows that India is set on the path of progressive Healthcare for each individual. The Indian (union and states) government spends 1.13% of the total asset from its current GDP [1]. According to the NITI Aayog report, India has inadequate and fragmented delivery of healthcare services, perhaps because of cultural and religious diversity or inadequate implementation of health policies.

According to a previous report [1], NITI Aayog has reported the following reasons for challenges, opportunities, benefits, and options for improving India's health sectors:

- ✓ **Strong economic foundation and policy implementation for transforming the healthcare industry, which is currently underperforming.**

The Indian economy is growing at a high rate. For a decade, India has controlled inflation, increased its GDP, and encouraged states to become policy-driven. The health factor of Indians has increased in the decade because of a solid economic background. The use of EHRs, medical health analysis, competent and expert health advice through machines, and data analysis through wearable devices is adopted in India. The "make in India" [2] policy helped investors invest more assets in the healthcare industry, resulting from the quality improvisation of health and healthcare industries in India.

Table 2 [3] summarizes Indian health systems with key performance indicators such as GDP, PPP, and global healthcare rank. The source was obtained from the Lancet journal.

Table 2. Key performance indicators and source: Lancet Journal.

	India	China	Sri Lanka	Indonesia	Egypt	Philippines
Total health expenditures as % GDP	4.0%	5.5%	4%	3%	5%	4%
Fiscal health expenditures as % GDP	0.9%	3.2%	2%	1%	1%	1.3%
Per-capita health expenditures (PPP)	239	761	491	363	516	342
Level of out-of-pocket (% Total health expenditures)	64%	36%	50%	60%	62%	54%
Neo-natal mortality 1980	60	65	24	41	53	27
Neo-natal mortality 2016	25	5	7	13	12	13
Global healthcare rank	145	92	71	138	111	124
The burden of disease (DALYs per 100,000 population)	34,000	26,300	24,000	28,900	28,000	31,000

- ✓ **Healthcare system improvisation can decrease mortality and poverty rates and accelerate economic growth.**

Implementing critical, intelligent, and automated health systems can help Indian people improve health and reduce mortality and poverty rates. The use of AI and machine learning algorithms in digital Indian healthcare data can enable the early prediction of diseases and provide remote-level advice from medical experts. Centralization of medical healthcare records can help rapidly access patient information and increase healthcare record utility.

- ✓ **Unnecessary and non-uniform health sector fragmentation is the main problem of healthcare industries in India.**

Data fragmentation at any level and its granularity are the significant reasons for the underperformance of India's healthcare systems. Fragmentation, a myriad of organizations, institutions (formal and informal rules), management, and administrative arrangements and entitlements that do not coordinate harmoniously and are often subjected to contradictory incentives, severely hampers the continuity of care and portability benefits [1]. In other countries, healthcare system fragmentation occurs with uniform policies and rules and has the same set of roles to access it. Uniform and limited fragmentation of the healthcare data and users helps:

- Protect the healthcare data from a third-party unauthorized entity;
- Have uniformity in centralized health systems.

Indian healthcare policies are growing and are adaptive and structured. Several problems of healthcare systems required to secure the systems are addressed.

4. Healthcare Industry and Applications in India

India is emerging in terms of revenue and employment in the healthcare field. The advances of ICT help the healthcare sector streamline data structure, access, and health analytics [14–16]. The healthcare sector in India is growing relatively slower due to its extensive coverage, strengthening services, and increasing expenditure by public and private players. In India, the healthcare industry is divided into public and private industries. The public healthcare industry (operated by the Indian government) is responsible for providing primary health services and treatments primarily to people in rural areas. The private sector provides amenities and services to the middle-class and upper-class people in India.

4.1. Segments of Indian Healthcare Industry

The Indian healthcare system is divided into six major segments: Hospitals, pharmaceuticals, diagnostics, medical equipment and supplies, medical insurance, and telemedicine,

according to the IBEF report [15]. The Infographics (Figure 5) describe each segment involved in the Indian healthcare system.

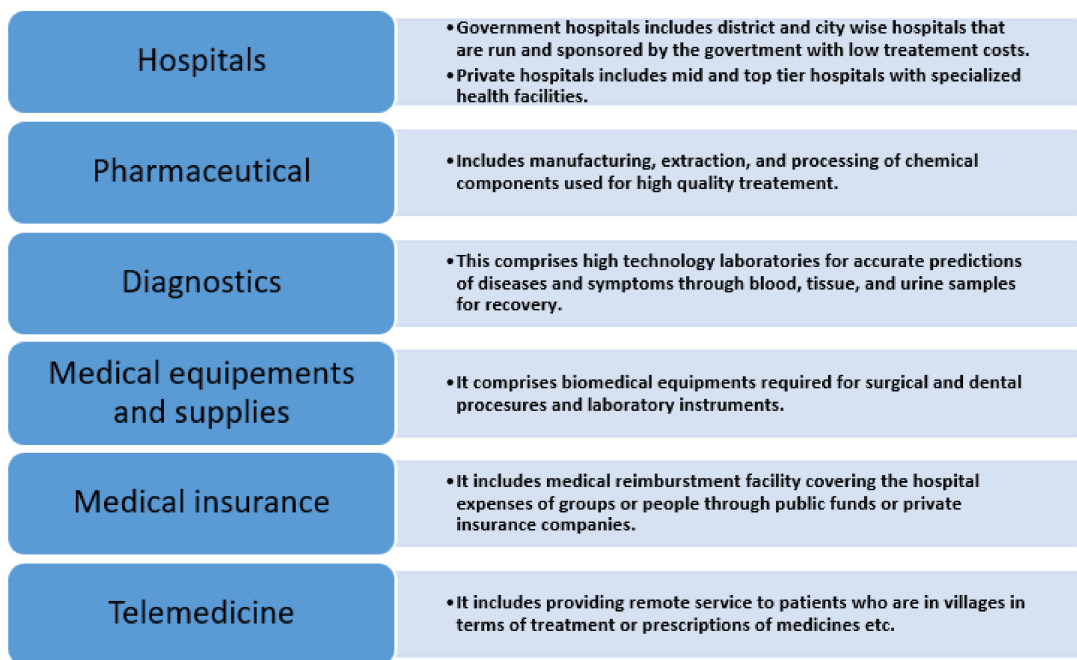


Figure 5. Segments of the Indian healthcare industry [15].

4.2. Advantages of the New Indian Healthcare Industry

Make in India is the fundamental initiative taken by the Indian government. Under this flagship campaign, the healthcare sector comprising hospitals, diagnostic centers, drugs and pharmaceuticals, and medical devices is identified as a part of the initiative [17,18]. Because of this initiative, the healthcare industry is transforming, and Figure 6 presents its benefits as infographics.

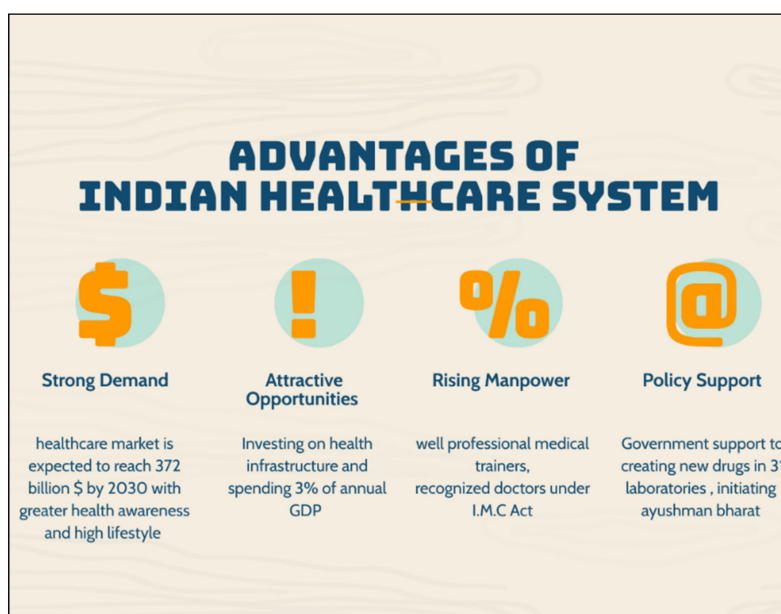


Figure 6. Advantages of the Indian healthcare industry [18].

4.3. Rise in Healthcare Infrastructure in India

With the advent of technology, the number of healthcare markets and assets in India has been increasing each year. India will have potential healthcare markets shortly. Many medical institutes are emerging because of the changes in government policies. The Indian government is motivating and encouraging medical colleges to be equipped. According to the national health profile in 2018–19 [17], the number of medical educational infrastructures in India has increased rapidly in the past 26 years. The total number of medical colleges in FY 2019 was 529, with 1,154,686 doctors with recognized medical qualifications. The presented demographics indicate increased medical colleges and doctors with recognized medical degrees in India (Figures 7 and 8).

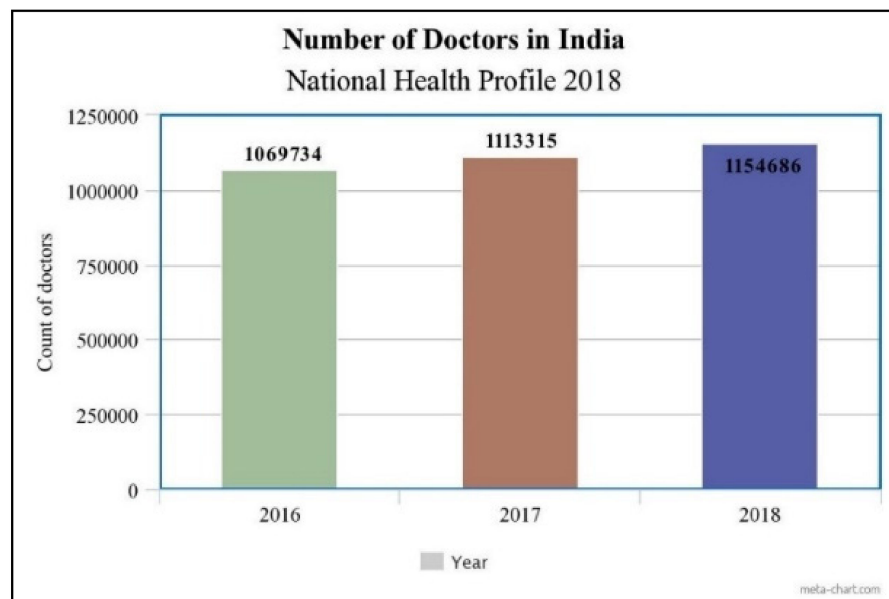


Figure 7. Number of doctors in India [17].

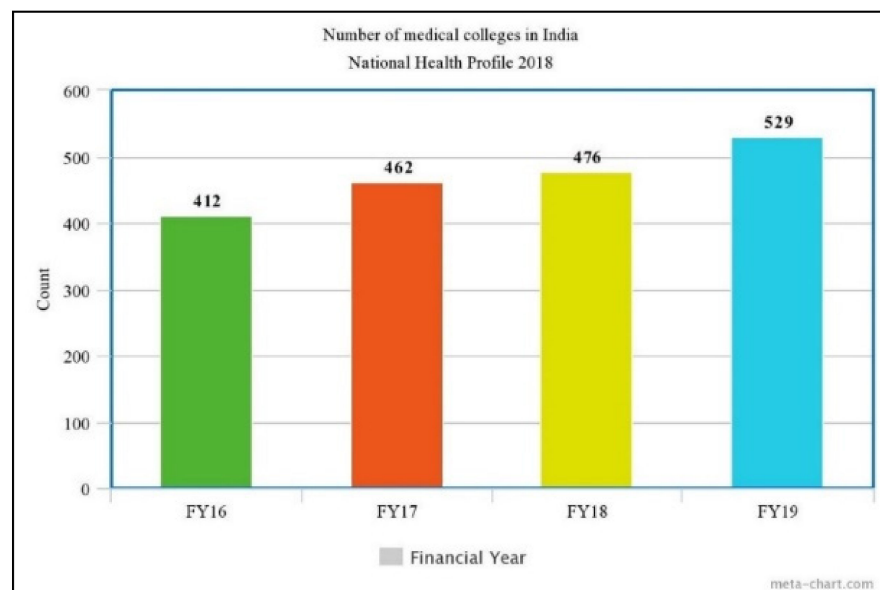


Figure 8. Number of medical colleges in India [17].

4.4. Trends in Indian Healthcare System

According to the article published in the Journal of Ayurveda and Integrative medicine [19] as well as the IRDA, CII, Grant Thornton, Gartner, and Technopunk [20], the notable trends observed in the Indian healthcare system are presented in Table 3. Each trend also has some privacy concerns and implications, which are listed in the same table (Table 3).

Table 3. Comprehensive overview of ‘trends and privacy implications in the Indian healthcare system.

Trends	Description	Privacy Implications
Community disease to personal- and lifestyle-related diseases	Due to urbanization and technology use in our daily lives, specific lifestyle-related diseases have led to community diseases. These diseases include cholesterol, blood pressure, diabetes, liver problems caused by overconsumption of products such as alcohol. This trend requires customized medicines and treatment with personal health and self-care.	In general, diseases are less harmful and can be treated by a local doctor or healthcare professionals. The record of such diseases and treatment is stored by individual hospitals in digital form. Local hospitals are not trustworthy enough to store such sensitive information, and hence it may invite data breaching issues due to insecure storage of EHR [20–22].
Healthcare expansion to Indian cities	The privatization of Healthcare by the government helped expand Healthcare to tier II and III cities. Hospitals are also built-in villages and rural areas to provide Healthcare to middle and lower-class people. The government is reducing taxes for the first five years for such businesses to encourage healthcare expansion in the private sector.	Hospitals built in rural areas mostly use paper-based prescription records; using these records lacks electricity and the Internet to run digital equipment. Paper-based records can be easily stolen and are most vulnerable to stealing personal information than digitally stored information.
Telemedicine	Many healthcare startups such as Apollo and AIIMS are adopting telemedicine services. Telemedicine can bridge the rural-urban gap to provide medical facilities, low-cost consultation, and diagnosis facilities to the remotest of areas through the high-speed Internet and telecommunication.	Research says that protection against privacy concerns is more in telemedicine, Which requires a multi-disciplinary and multi-stakeholder approach. Most of the employees working in the telemedicine industry are either very busy with workload or are less trained about security and privacy violations. This invites data breaching attacks, phishing attacks, unauthorized access, and so on [23].
AI in Healthcare	The adoption and use of AI-based healthcare applications are rapidly growing. AI helps solve the problems of patients, doctors, hospitals, and the overall healthcare industry.	AI ensures that the disease prediction is way faster than an actual spread of disease due to prediction technology. However, with the massive use of data for training purpose (which contains sensitive data), privacy issues of identifying individual rises [24].
Home healthcare [24]	Home healthcare provides Healthcare at affordable prices at patients’ homes. It saves the traveling costs of doctors/patients, and treatment is provided with minimum logistic interventions.	-

Table 3. Cont.

Trends	Description	Privacy Implications
Growth of health insurance	Health insurance is gaining momentum in India. The trust of people in India in health insurance and the assurance by health insurance has increased in past years. Many companies such as Aditya Birla and LIC provide health insurance to people.	There are massive amounts of data generated in the health insurance and needs to be protected. The sensitive information such as personal details, disease details, past health history is being recorded and shared with the third party by insurance companies without individual consent. Most of the data shared with the third party are non-anonymous; it is straightforward to predict the identity of individual human beings.
Mobile-based health delivery	The mobile solid technology infrastructure and the launch of 4G can drive mobile-based health initiatives in the country. It enables fast health-related services with reduced costs and superior reach [21,22].	-
Technology for health	Technological intervention is increasing in India. According to [21], India's medical technology sector can reach US \$9.60 billion by 2022. Various advent technologies are used in the healthcare domain, such as machine learning algorithms for prediction of specific health parameters/data/diseases/behavior and the Internet of things-based healthcare systems [22]	-
Luxurious living and health	Luxurious services, including pick and drop facilities, doctor visits at homes, online prescriptions, have become a part of the Indian healthcare industry.	-

4.5. Popular Healthcare Hubs/Startups in India

The competitive value of the Indian healthcare system provided well-trained medical professionals, well-equipped instruments, ubiquitous healthcare applications, and suitable patient-centric health hubs. Table 4 presents some healthcare hubs.

Table 4. Healthcare hubs in India.

Health Care Providers	Statistics	Description
Apollo Hospitals	9844 beds 70 hospitals 8500+ doctors Total income was Rs 9648.88 crore (US \$1.38 billion) in FY 19 and Rs 8347.39 crore (US \$1.19 billion) during 20	Apollo healthcare has hospitals and pharmacies across India. Moreover, the company provides project consultancy services, health insurance services, education and training programs, and research services. It also operates birthing centers, day surgery centers, and dental clinics.
Thyrocare Technologies Limited	Over 1200 employees 571 cities Consolidated total income of Rs 412.86 crore (US \$59.07 million) in FY 19 and Rs 337.43 crore (US \$48.28 million) in 9 MFY 20	Thyrocare is the first completely automated diagnostics laboratory. The company offers cancer and HIV diagnostics centers, chemotherapy, and dialysis centers across India.

Table 4. Cont.

Health Care Providers	Statistics	Description
Fortis Healthcare	36 healthcare facilities Approximately 9000 beds 415 diagnostic centers Total consolidated revenue of Rs 4469.35 crore (US \$639.48 million) in FY 19 and Rs 2379.79 crore (US \$340.51 million) in H1FY 20.	Fortis healthcare is considered an integrated healthcare delivery service provider in India. Fortis memorial research institute (FMRI) ranked second in a study of 30 most technologically advanced hospitals in the world conducted by topmastersinhealthcare.com. 19 August 2020
Net meds	14 logistic centers across the country 24 × seven online portal and mobile application Three million downloads (till 2018) with more than \$512 million profit; it is projected to earn \$3.645 billion by 2022.	Netmeds is an online platform for the pharmacy industry. It offers significant pharmacy products through online shipments. It is also called “India ki pharmacy.” The mobile application is well-equipped with voice chats, e-mailing services, and 24 × seven customer care services.
Practo	Free services for doctors and patients Focused website and 1 lac doctor profiles in India	Practo is the world’s leading healthcare platform and works as an independent medical portal, connecting doctors and hospitals across India and the globe. Although the reach of Practo is global, it was founded in 2008 in Bangalore. Practo provides its users with diagnostic search features on its web-based platform through high-quality photographs and filter options. Practo is suitable for private doctors who independently run hospitals in rural and urban areas.

5. Healthcare Data Utility: Context of India

In recent years, healthcare analytics has become a broad research topic in India. Automation in healthcare processes is crucial because the healthcare system has transformed from diseases- to a patient-centric system. Many people suffer from diseases related to their lifestyle, eating habits, and work profiles. Such diseases include diabetes. India’s overall healthcare data system is being transferred from offline to online, wherein digitization of the health data such as EHR, healthcare applications, digital X-rays, and reports is being used for diagnosis and treatments. A large amount of data is generated on server-side platforms. Such healthcare data must be analyzed for suitable treatments and early disease predictions. In the healthcare industry, technologies such as data analysis, machine learning, AI, and blockchain [23,24] play a vital role because they enable healthcare systems to systematically use and analyze the data to identify inefficiencies by keeping them secure and provide optimal practices that improve care and reduce costs. Doctors, clinicians, healthcare researchers, and medical industry professionals are the beneficiaries of healthcare analytics. This section answers the following questions [24]:

- What are the different data sources of the Indian healthcare industry?
- How are the Indian healthcare data classified?
- What are the different intelligent platforms used for healthcare data analysis?
- Which technologies drive the Indian healthcare system?

5.1. Sources of Indian Healthcare Data

The comprehensive data can be divided into three types in India, namely clinical, exogenous, and genomic data.

- **Clinical Data**

The clinical data include electronic health records such as magnetic resonance imaging, X-rays, blood and urine record, molecular imaging, ultrasound, photoacoustic imaging, and fluoroscopy data. Moreover, it comprises pharmaceutical records and sociodemographic details of patients or populations. The clinical data are collected from health institutes, hospitals, or health insurance companies through interviews, surveys, or patient treatment. Such data contain compassionate information, including the diseases, medicine prescribed, patient deficiency, address, and other personally identifiable information. Such data can be structured or unstructured, depending on the type of data collected from patients.

- **Exogenous Data**

The data obtained from medical and wearable devices such as pacemakers, fitness bands, and intelligent watches are categorized as exogenous data. Because data are collected from devices and sensors, they are stored in the cloud and are considered to have sensitive information such as heart rates and blood-related information; such data are breaches to privacy issues [25,26]; the details are presented in the next section of the paper.

- **Genomic Data**

The data relating to people's genes and genetic structures are classified as genomic data. The data are susceptible but simultaneously highly complex to understand as far as their analysis is concerned. The genomic data analysis helps scientists and medical practitioners to predict remedies accurately and preventive methodologies for patients with a particular lifestyle, that have similar genetic make-up, or are exposed to similar environmental conditions [27,28].

5.2. Types of Healthcare Data

The main three types of data [29] are structured, unstructured, and semi-structured data. Table 5 presents the details of these data types.

Table 5. Type of healthcare data.

Type	Description	Examples	Clinical Data	Exogenous Data	Genetic Data
Structured	The data arranged in a structured format are considered structured data. Such data are primarily arranged in rows and columns. The structured data are mostly easy for analysis but have considerable sensitive information and direct identifiers.	Blood reports, sugar reports, billing information of patients, and Indian census information. Primarily clinical data can be categorized as structured data.	Yes	No	Yes
Semi-structured	The data that is minimally structured and requires scripts for extraction are classified as semi-structured data. Such data are generally captured from wearable devices, which monitor a person's response to particular medicine and activity.	XML-, JSON-extracted reports are considered semi-structured. The exogenous data are generally considered semi-structured data.	No	Yes	Yes
Unstructured	The data with no uniform format or structure are classified as unstructured data. Human written prescriptions and reports are considered unstructured data.	The doctor's written prescription on a notepad, images, videos, or time series reports is considered unstructured.	Yes	Yes	Yes

5.3. Key Data Sources of Health Information of India

In India, health is a state subject; the constitution places the responsibility of health on both central and state Governments. In various data sources, the health information of patients is preserved. The central government is responsible for provisions listed in the union list, and the state government is responsible for providing medical services and amenities such as hospitals and dispensaries [30]. The data sources are divided into direct and indirect sources. Table 6 presents the features, advantages, and limitations of these critical sources of health information.

Table 6. Sources of the Indian healthcare system.

Data type	Direct/Indirect	Description	Strengths	Limitations
Population census	Indirect	Population census is about storing information on the population of India. The vast database comprises social, geographical, and demographic information of people collected every ten years.	It covers small information for people across India. In terms of analysis, the data are most helpful with proper predictions of literacy rates, social-cultural activities, and food consumption habits.	Population census is conducted every ten years by the government of India. The health-related information is considerably low. Providing health analytics on census datasets is challenging.
Civil registration system	Indirect	The count and details of the population are obtained in certain situations, such as the birth and death of a person and from lost and found records.	Primarily the data related to time and location are recorded.	Death-related information is documented correctly. For example, minor details such as the cause of death (crime, health issue, and aging), city information, the place of death (can be different from the place of living) are not correctly stored, and hence the data are not productive for analysis. Inadequate information or columns are generally filled by some random values, which cause wrong/improper prediction of specific facts.
Public surveys	Direct	The data generated by companies, work organizations, health insurance companies, and third parties are considered the survey data.	The data are primarily structured for a particular party collecting data. The data are collected through digital mediums, including online forms, SAP systems, web portals, and social media.	The irregular and impure data are obtained most of the time. Mostly, the data are not accurate because not all people provide correct information. The data quality is low. The data remain moderately analytical.
Service records	Direct	The data captured directly from people about their health, work, and other demographic aspect. Work organizations generally capture the data, and data use is limited to organization scopes and limits. The organization must acquire consent from its employees for sharing, using, and publishing such data on a public platform or with a third-party service provider.	This is exclusively used for service management. The data quality is excellent and reliable. The data can be captured in limited time intervals of months or days (depending on data captured).	Mainly, data duplication and inconsistency problems arise because the data are not always at the centralized place.
Administrative records	Direct	It contains information regarding family details, financial planning, personality, and emotional details.	The data are captured as a single data source and generally are in a suitable quality format. These data are highly analytical but are generally substantially private.	The privacy problems and data breaching probability are maximum. The data require uniform and secure policies to access, use, share and publish.

5.4. Technologies Used in Healthcare Data: Indian Perspective

5.4.1. Data Mining / Analysis of Indian Health Data

Data mining is an essential and promising technique used in most countries for healthcare data analysis. Most data are decentralized and maintained by laboratories, medical centers, and hospitals in public and private forms in India. Data mining techniques such as association, classification, and clustering are used by healthcare [31–34]. The details of data mining techniques are as follows:

- **Clustering:** Clustering is a type of unsupervised learning and is slightly different from classification. In clustering, many datasets are divided into small chunks (clusters) based on some similarity. The Euclidian distance is used to calculate the relation/distance between two data values. K-means clustering is a prevalent method of clustering; however, it is time-consuming and slow.
- **Classification:** Classification comprises training and testing data. Training is required as it helps create classification rules. For high accuracy, providing maximum data for training is an optimal practice. The accuracy of a classification model depends on the degree of classifying rules being true, which is estimated using the test data.
- **Association:** Association is an exciting mining method in which frequent and usual patterns in a dataset are determined. It is also known as the market basket analysis because it can identify the association among purchased items or unknown customer sales patterns in a transaction database. Association rule mining is widely used

for identifying the relationships among various symptoms with similar causes of particular diseases.

- **Regression:** Regression is used to find the correlation among various attributes defined over a particular function. For regression, a mathematical model is constructed with the training data with dependent and independent variables. Regression can be linear and nonlinear. Linear regression identifies the relationship between a dependent variable and one or more independent variables. Logistic regression, a nonlinear regression type, can accept the categorical data and predict the probability through the logit function.

5.4.2. Artificial Intelligence, Machine Learning, and Deep Learning in Indian Health Data

Big data plays a vital role in Healthcare, bioinformatics, and health informatics in recent years [35–38]. The comprehensive data generated in healthcare in 2009 are 44 times higher than that generated now, and sometimes, processing the data for analytics, prediction, and accurate visualization is difficult. AI has gradually been evolving in healthcare. According to an article in Wired, in India's Aravind eye care system, ophthalmologists and computer researchers work together to test and deploy an automated image classification system to screen millions of retinal photographs of diabetic patients [39–41]. The technologies including AI and deep learning have potential applications in the healthcare industry such as medical imaging, radiology, cancer prediction, diabetes and heart disease prediction, pathology, genome interpretation, and patient monitoring.

Moreover, AI is used in fraud detection in the healthcare field. Fraud can be committed by service providers or healthcare insurance companies in billing and treatment or other report generation [42].

5.4.3. Data Visualization of Indian Health Data

Interactive data visualization helps understand many data, particularly in decision support systems [43–45]. The volume of the data collected in the healthcare domain is large and unstructured and increases rapidly. The high-dimensional data are sensitive and meaningful for health analytics and prediction. Some healthcare information (particularly about diseases such as COVID 19 and swine flu) varies with time; hence, their effect can be analyzed with a time parameter. Regarding visualization, empirical experiments show that visualizations facilitate an understanding of clinical data; however, a consistent method to assess its effectiveness remains unavailable. Focusing on the Indian data collected manually and digitally, representing the data on a large platform (e.g., dashboard and data publishing platforms) is considerably complex. Many datasets have privacy concerns, and hence their visualization is affected and distorted from actual outcomes.

5.4.4. Augmented and Virtual Reality in Indian Health Data

For addressing the big data present in the healthcare industry, visualizing the complex data more simply is crucial. The optimal visualization facilitates the understanding of health problems for accurate diagnosis and early predictions. The human perception of visualization is limited and requires high cost and time on the screen. Virtual reality is used in healthcare in many fields, such as psychological theory, cancer detection, and maxillofacial surgery.

Table 7 presents the details of studies and innovations based on various technologies in the Indian healthcare industry and the studies conducted on the Indian healthcare data, mainly from 2016. Over three years, 31 research papers have been published on disease prediction/diagnosis, which uses the Indian health data in principle.

Table 7. Existing research in the Indian healthcare industry.

Types of Disease	Type of Data	Data Mining Technique	References
Data Mining of Indian Health Data			
Conventional pathology data	Structured	Support vector machine classification	[32–34]
Heart disease	Structured and unstructured	Naïve Bayes, decision tree, and K-nearest neighbor	[35]
Lymphoma disease and lung cancer	Unstructured (image dataset)	Support vector machine	[36,37]
Psychiatric diseases	Structured and semi-structured	Random forest, support vector machine (SVM), K-nearest neighbor	[38,39]
Liver diseases	Structured	k-means (KM) clustering algorithm, agglomerative nesting (AGNES), clustering algorithm, density-based spatial clustering of applications with a noise clustering algorithm, ordering points to identify the clustering structure, clustering algorithm, and exception maximization clustering algorithm	[40–42]
Skin disease	No paper found on the Indian dataset		[43]
Diabetes	Structured	Improved K-means algorithm and logistic regression	[44]
Chest disease	Unstructured (image dataset)	Risk factor identification through correlation-based feature subset selection with particle swarm optimization search method and K-means clustering algorithms. Supervised learning algorithms such as multilayer perceptron, multinomial logistic regression, fuzzy unordered rule induction algorithm, and C4.5 classification algorithm	[45]
Chronic disease	Structured	Naïve Bayes, K-nearest neighbor, and decision tree	[46,47]
Breast cancer	No paper found on the Indian dataset		[48]
Cardiovascular diseases	No paper found on the Indian dataset		[49]
Parkinson disease	No paper found on the Indian dataset		[50]
AI/machine learning/deep learning in the Indian health data			
Conventional pathology data	Unstructured (image data)	Convolutional neural network on pathological myopia disease for vision blindness. Classification-based glottal closure instants detection from pathological acoustic speech signals	[51–56]
Heart disease	Structured	Naïve Bayes. Decision tree and random forest	[57]
Lymphoma disease and lung cancer	No paper found on the Indian dataset		
Psychiatric diseases	No paper was found on the Indian dataset.		
Liver diseases	Structured	Multilayer perceptron neural network algorithm based on various decision trees algorithms such as See5 (C5.0), chi-square automatic interaction detector, and classification and regression tree with boosting technique	[58]
Skin disease	Unstructured (image dataset)	Deep learning algorithms: (inception_v3, MobileNet, resnet, exception)	[59]
Diabetes	Structured	Linear kernel SVM, radial basis function kernel SVM, k-nearest neighbor, artificial neural network, and multifactor dimensionality reduction	[60]
Chest disease	Unstructured (image dataset)	Random forest	[61]
Chronic disease	No paper found on the Indian dataset		
Breast cancer	Unstructured	Thermolytic risk score framework	[62–64]
Cardiovascular diseases	No paper found on the Indian dataset		-
Parkinson disease	No paper was found on the Indian dataset.		-
Data visualization of the Indian health data			
Covid 19 dataset (open research dataset of India)	Structured	-	[65–67]
Covid 19 dataset (open research dataset of India)	Unstructured (X-ray image dataset)	-	[68–70]
Tuberculosis screening	Unstructured (X-ray image dataset)	-	[71–73]
Augmented and virtual reality in Indian Healthcare			
Automated data capturing from medical devices	Unstructured	-	[74]

6. Healthcare Data Privacy in India

Patient privacy is the most crucial aspect and jurisdiction of all countries worldwide, and all countries have accepted that the privacy of the people must be respected under any consequences. Privacy is a fundamental right of humans [75–77]. Some countries (primarily Europe and the USA) strictly prioritize privacy policies. Famous laws such as HIPPA [78] and GDPR [79] provide strength to people about their privacy concerns and help them build trust. However, to date, no universal definition of privacy is available. Some definitions are perception-centric and change with countries [80]. Various definitions of privacy are as follows.

Definition 1. *“The state of being alone or the right to keep one’s matters and relationships secret”—definition obtained from the Cambridge dictionary [81].*

Definition 2. *“No one shall be subjected to arbitrary interference of their privacy, family, home, or correspondence or attacks upon their honor and reputation. Everyone has the right to the protection of the law against such interference or attacks”—article 12 universal declaration of human rights.*

Definition 3. *“privacy can be divided into several separate, but related, concepts:”*

- *Information privacy involves establishing rules governing collecting and handling personal data such as credit information and medical and government records. It is also known as data protection;*
- *Bodily privacy concerns the protection of people’s physical selves against invasive procedures such as genetic tests, drug testing, and cavity searches;*
- *Privacy of communications covers the security and privacy of mail, telephones, e-mails, and other forms of communication; and*
- *Territorial privacy concerns the setting of limits on intrusion into the domestic and other environments such as the workplace or public space. This includes searches, video surveillance, and ID checks—Australian law reform commission.*

Definition 4. *“Privacy is the right to be let alone or freedom from interference or intrusion. Information privacy is the right to have some control over how your personal information is collected and used”—the international association of privacy professionals.*

Privacy plays a vital role in the healthcare field because healthcare data contain sensitive information about patients and related stakeholders. As stated in the introduction, most healthcare fields have changed their operations from disease to patient-centric. The records of people are stored according to their characteristics, personal and emotional behavior, and geographic information. Central storage of electronic medical data across highly configured servers is one of the most suitable options for data analysis [82,83]. It prevents the use of duplicated or redundant data because of its central storage and maintenance. The amount of healthcare data is significant, and in India, the data are also unstructured. Furthermore, this can lead to privacy and data breaching because of its storage- and transformation-related concerns.

In India, Privacy is not treated as a serious concern. Privacy issues in healthcare, specific to India, are caused by prevalent complacency, culture, politics, budget limitations, large population, and infrastructures. Due to these factors, data security requires a backseat that allows easy access to confidential information. Furthermore, the prevalent culture affects healthcare disclosure in India. In many cultures, disclosing sensitive personal healthcare data is considered ill-mannered. This leads to discrepancies in the recorded healthcare data and a decrease in the level of treatment meted out. The results and statistics of treatments given do not match the records due to inaccurate data reporting.

India is a democratic country with a large population, and maintaining a standard infrastructure is a problem for implementing privacy models in India. The cost required to implement a privacy model is substantial and requires funding from the government

and people. Making the privacy model a success involves the work of specialists in the Privacy and healthcare fields. Budget constraints may cause an ineffective model to be implemented, which cannot be secure and safe from attacks.

According to the recent news, the Indian health ministry has proposed a law to govern data security (personal data protection bill) that would provide people complete ownership of their data. People can access, share, and deny sharing the records available at the server. The health ministry proposed digital information security in the healthcare act on March 11, 2018. The committee suggested the following key points and developed a privacy framework:

- The law must be flexible and adhere to changing technologies.
- Law must be applied to public and private sector entities.
- Entities controlling the data should be accountable for data processing.
- Consent must be structured and genuine.
- Data processing and analysis must be minimal.
- A high-powered statutory authority should enforce the data protection framework.

The Indian healthcare data are considerably diverse and collected from different heterogeneous sources (public and private sector hospitals and health insurance). No regulations are enforced over the health data authorship, due to which any third party can access the sensitive data and misuse the data for its benefit. The proposed law has guidelines and technological aspects for preserving healthcare data privacy.

Privacy Issues in the Indian Healthcare System

From the literature survey, this paper presents 14 privacy issues (10 Primary issues) specific to the Indian healthcare system. Figure 9 illustrates the privacy problems. Each privacy issue is described with an appropriate example obtained from the available resources [84].

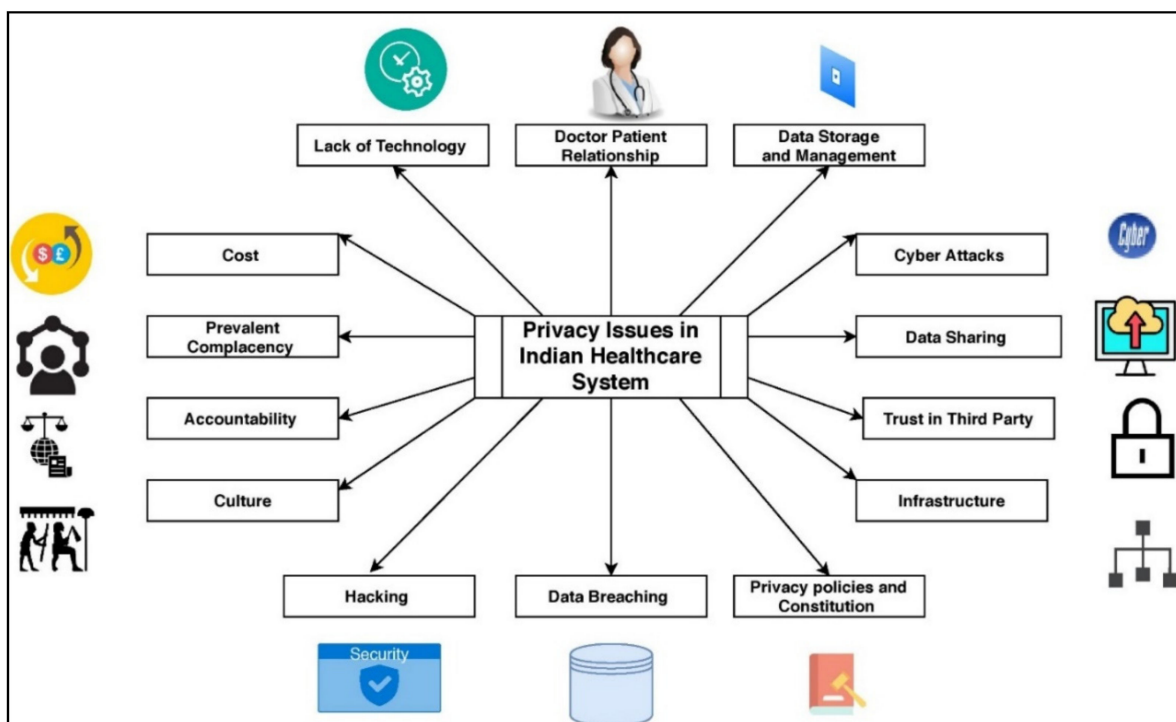


Figure 9. Privacy issues in the Indian Healthcare system.

A. Lack of Technology and Infrastructure

Healthcare technology is changing, and the paper-based records of the patient are not used anymore. The records are being converted into EHRs for easy digital access through the Internet [85]. The use of wearable technologies, patient monitoring through sensor networks, and data analysis of patient records for early disease prediction are being implemented to ease and improve the lifestyle. Despite numerous advancements in this field, India has not successfully implemented technologies at the ground level of the healthcare system. According to the Indian healthcare system (Figure 4), the village/taluka and district hospitals still lack technologies. Another viewpoint is that patient records are primarily stored using paper-based technology rather than centralized electronic technology. In India, >60% of the area are villages, which indicates that the healthcare sector is majorly based in villages. Currently, rural areas face problems such as lack of electricity, high-speed Internet, and high-technology medical equipment facilities in hospitals and dispensaries.

The lack of technological interventions directly affects patient privacy. Most villagers must physically visit healthcare centers and hospitals for their treatment with prescriptions, and health advice is provided on papers. Because records are paper-based, there is no control over who is having access at what level. Consent management, storage guidelines, and access control are no longer applicable to paper-based medical records.

B. Doctor–Patient Relationship

Another threat to Privacy in India is the trust between doctors and patients. Most people from tier II and III cities do not trust their doctors and hospital staff for their data security. Most hospitals share data with a third party without acquiring patient consent [85,86]. No law covers such actions because no uniform policy regarding such fraud is defined in the constitution. According to the literature, many factors lead to the doctor–patient relationship being compromised.

1. Poor government health systems.
2. A poor ratio of doctors to patients.
3. Easy accessibility of information and privacy concerns [86].
4. Lack of a role of the patient in the decision-making process.
5. Corruption [87].

C. Data Storage and Management

With the large population of India, storing electronic medical records in the cloud is crucial. A survey [88] revealed that most Indians store their health records on the cloud for easy access; however, considerably few are concerned about their privacy. Most Indians store sensitive data on the cloud, relying on the fact that the cloud provides security. Additionally, data management makes design most optimal for information, centralized or distributed storage, and data confidentiality and demand.

D. Cyber Attacks and Hacking

A dynamic EHR is maintained and regulated by a third party for suitable data storage and management. When such data are shared either for analysis, research, or marketing purposes, maintaining patient privacy is the responsibility of the third party. Moreover, the third party is responsible for providing sufficient mechanism security for data storage to prevent cyberattacks. Most Indian healthcare data are either stored in outdated systems, or no security mechanism is applied to the data. A suitable and secure system must not allow access to the data to unauthorized users. Table 8 presents cyberattacks presented in the Indian data [89].

Table 8. Cyberattacks presented in the Indian healthcare data.

Cyber Attacks on Data Gathering Phase	Cyber Attacks at Network Phase	Cyber Attacks at Storage Phase
Phishing attack	Eavesdropping of health record	Cross-site scripting attack
Log access attack	Man-in-the-middle attack	Weak authentication attack
Social engineering attack	Data tampering	
	Denial of the service attack	
Brute force attack (on passwords)	Data interception	SQL injection attack.
	Spoofing and sniffing attack	

E. Data Sharing Trust in the Third Party

Most Indian healthcare systems lack consent; no consent is acquired while sharing data with a third party. Most private organizations share the data of their employees with a third party without applying any rules or policies to it. Sharing data for research and analysis purposes presents no harm; however, such information is shared with personal identification information. Alternatively, the information is the responsibility of organizations, which are not regulated by any statutory body for misuse. Most public sector hospitals share their data without patient consent [90].

Sharing data with a third party always presents doubts in the trust parameter. The privacy model is divided into two types: Trusted and untrusted models. In the trusted model, the data owner trusts the third party and shares the health data, and an untrusted party does not gain the owner's trust. The use of a third party questions the confidentiality and integrity of the data and makes dealing with the party during the development of a highly reliable health architecture a vital issue.

F. Lack of Policy and Constitutional Limitations

Privacy is always ignored; private organizations especially give less importance to privacy. In Indian healthcare systems, privacy policies remain inadequate. Several case studies on the Indian healthcare domain have indicated that we are far from privacy implementation and applying privacy rules through design principles. A cohesive privacy policy must be implemented in India. The following questions remain unanswered and must be addressed when the data are used, shared, and published by a third party or organization [91]:

1. Who owns and accesses patient records and why?
2. What type of data with what granularity level must be collected?
3. Where must the data be stored (central warehouse or hospital)?
4. Who can view medical records?
5. Who is responsible for disclosing medical records?
6. Which consent must be acquired while deleting patient records?

G. Data Breaching

Despite taking data security measures, data breaching is one of the significant privacy problems in India. In healthcare, the main reasons behind data breaching are as follows [90]:

1. Broken access and authentication.
2. Flawed service level agreements by organizations.
3. Poor backup and recovery plans in case of data loss.
4. Reverse engineering methods.

A cross-sectional data sharing system is the most suitable technique for low- and middle-income countries like India. India's Aadhaar personal identification program is promising. It is responsible for generating and monitoring health and social data, including EHRs, through a unique identification number. A unique card is distributed to all Indian citizens for identification. In 2017, the supreme court of India addressed privacy concerns,

including breaching the Indian health data stating that because the Aadhaar card contains sensitive information, it cannot be used as a mandatory document in fields such as banking, the insurance sector, and mobile servicing [92].

H. Culture

Cultural interventions are other challenges faced by the Indian healthcare industry. Most cultural communities do not allow people to share or disclose their personal information due to predefined cultural restrictions, resulting in the recording of false information. The discrepancy in health records results in inaccurate analyses and an inaccurate treatment that is meted out [93].

I. Prevalent Complacency

Complacency is widespread in Indian healthcare. A large amount of work, planning, cooperation, and communication among multiple departments is required to make the privacy and security of healthcare in India a success. However, due to slackness, the probability of the privacy model implemented in India is poor [94].

J. Cost

Implementation of the privacy model with suitable infrastructure is costly. The Indian government faces other problems such as poverty and corruption that are given high priority, and privacy model implementation is given the least priority. Small organizations do not have enough assets to protect their employees.

Table 9 presents the privacy problems of the healthcare system of India as mentioned above. Table 10 explains how each privacy problem is being addressed in different healthcare structures in India.

Table 9. Privacy issues stated in the Indian healthcare system.

Type of Healthcare	Privacy Issues Stated in the Indian Healthcare System.						
	Type of Sector	Lack of Technology	Doctor-Patient Relationship	Data Storage and Management	Cyberattacks	Data Sharing	Trust in the Third Party
Super specialty hospital	Public	Very good	Trustworthy	Very good	Minimal risk	Drafted Policies for consent	Strict
Medical institutes/colleges	Public	Adequate	Trustworthy	Good	Minimal risk	With good consent	Easy
District and taluka hospitals	Public	Adequate	Bit trustworthy	Adequate	High risk	Minimal consent	Easy
Primary healthcare centers	Public	Poor	Bit trustworthy	Poor	High risk	Without consent	Easy
Village hospitals	Public	Poor	Not trustworthy	Poor	High risk	Without consent	Easy
Super and multispecialty hospitals	Private	Excellent	Most Trustworthy	Very good	Minimal risk	Drafted Policies for consent	Strict
Tier II and III city hospitals	Private	Very good	Most Trustworthy	Very good	Minimal risk	Drafted Policies for consent	Strict
Private clinics	Private	Very good	Trustworthy	Very good	Minimal risk	Drafted Policies for consent	Strict
Non-profit organizations	Private	good	Trustworthy	Very good	Minimal risk	Minimal consent	Strict
Pharmaceutical industry	Private	Excellent	-	Very good	Minimal risk	Drafted Policies for consent	Strict
Health insurance company	Private	Excellent	-	Excellent	Minimal risk	Drafted Policies for consent	Strict
Private organizations	Private	Excellent	-	Excellent	Minimal risk	Minimal Consent	Easy

Table 10. Privacy issues stated in the Indian healthcare system (continued).

Type of Healthcare	Privacy Issues Stated in the Indian Healthcare System.							
	Type	Infrastructure	Privacy Policies	Data Breaching	Hacking	Accountability	Cultural Interventions	Cost
Super specialty hospital	Public	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	Average
Medical institutes/colleges	Public	Adequate	Well defined and followed	Adequate probability	Adequate probability	Yes	No	Low
District and taluka hospitals	Public	Poor	Well defined but not followed	High probability	High probability	No	Yes	considerably low/no cost
Primary healthcare centers	Public	Considerably poor	Well defined but not followed	High probability	High probability	No	Yes	considerably low/no cost
Village hospitals	Public	Considerably poor	Not defined	High probability	High probability	No	Yes	No cost
Super and multispecialty hospitals	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	Very high
Tier II and III city hospitals	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	High/considerably high
Private clinics	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	High
Non-profit organizations	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	Low
Pharmaceutical industry	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	NA
Health insurance company	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	NA
Private organizations	Private	Excellent	Well defined and followed	Less probability	Less probability	Yes	No	NA

7. Open Issues and Further Discussions

With the emerging healthcare sector from a revolutionary perspective, India is growing in healthcare analytics rapidly [95–98]. Healthcare 3.0 was patient-centric from a disease-centric approach. However, healthcare 4.0 uses various technologies like IoT, Blockchain, Machine learning, and AI to identify and predict disease, analyze historical health data, and other intelligent healthcare applications. Such newer technologies require extensive data and fastest accessing resources, which ultimately need equivalent security techniques to protect them [99]. In India, the security and privacy of healthcare data are always complex and challenging tasks. The reasons are already discussed in the previous sections. Based on the detailed overview of the Indian healthcare system and its privacy issues, there are some open issues and further discussions elaborated in this section.

Privacy issues in healthcare-specific to India are due to prevalent complacency, culture, politics, budget limitations, huge population, and infrastructure. Due to these factors, data security takes a backseat allowing for easy access to confidential information.

The prevalent culture also affects healthcare disclosure in India. In many cultures, the disclosure of sensitive personal healthcare data is looked down upon. This leads to discrepancies in the healthcare data recorded and a decrease in the level of treatment meted out. Research and statistics of treatment given then do not match the records due to inaccurate reporting of data.

India is a country of large democracy and large populations; maintaining standard infrastructure is another issue of implementing privacy models. The cost required to implement a privacy model is substantial and requires funding from the government and individuals. To ensure the privacy model is a success, it involves specialists in privacy and the field of healthcare. Budget constraints may lead to an ineffective model being implemented, which will not be secure and safe from attacks.

7.1. Key Performance Indicators in the Context of Privacy in the Indian Healthcare System

Key Performance Indicators (henceforth termed as a KPI) are very important as they measure privacy concerns that need to be addressed first. According to the research done in [100,101], the following KPI is related to privacy issues; however, they are applicable for social media photos and video sharing in the existing research. Since privacy issues exist everywhere, in every field, KPI is applicable in the Indian healthcare context as well. The list of KPIs and their details are given below:

- **Forced Trust vs. Control:** A forced trust is a trust in which an individual has no choice but to trust any healthcare system. On the other hand, control is a systematic view of obtaining trust and assuring each individual that their sensitive personal data will not be shared with the third party without any consent. In the Indian healthcare context, the ratio of forced trust to control is high. People tend to have less trust in any healthcare system because of constitutional limitations. This is one of the significant KPIs in the context of Indian Healthcare privacy.
- **Content Viewed by Whom:** Though there is limited access to any EHR and only authenticated people can view or access the sensitive data, there is still the possibility that unauthorized entities may access health records. Weak passwords, inappropriate security policies, conflict in access controls, and sharing passwords to untrusted persons are the possible reasons sensitive data may be misused. In India, the healthcare system is not very structured and centralized. Local hospitals keep their records on local servers, which are highly vulnerable to various attacks. Data breaching primarily happens in tier 2 and tier 3 cities and village hospitals.
- **Tacit Knowledge:** Even if the healthcare system ensures maximum protection against data breaches for healthcare data, the metadata or tacit information may reveal more information than basic health information. Using reverse engineering techniques or social media analysis, it is easy to gain personal information. In India, there are many cases reported against criminals who seek sensitive information through social media accounts. Unfortunately, there is no control over the protection against such information.
- **Laws and Regulation:** Limited regulation and law in the constitution are essential KPI in the Indian healthcare field. As per the latest data of 2019, Indian healthcare generates 1021 gigabytes of data per year. Managing such a massive amount of data by protecting sensitive content must prioritize the Indian government.
- **Use of new Data Protection Technologies:** The newer technologies like blockchain, two-factor authentication, machine learning, AI, and attribute-based anonymization is only implemented in high-end industries or healthcare organizations. Small sector health organizations, village, or tier 3 hospitals do not have funds to support the protection of such data, and hence newer technologies cannot be used.
- **Researcher's Satisfaction:** Since there is a massive generation of healthcare data, it is an excellent opportunity to analyze the data for research purposes. Healthcare analytics is the emerging field of computing and is rising exponentially in India. More restricted and policy-imposed data are not suitable for analytics purposes, and data quality gets degraded.
- **Industry-academia collaboration exists for privacy preservation mechanisms:** There is a huge gap between industry and academia in India. Despite having good researchers in the privacy field, their work is not reaching the industry.

It is also noted that all these KPI are not treated with the same priority. In order to priorities KPIs, the different stakeholders are taken into consideration. This paper's contribution is to summarize the ratings given by stakeholders directly or indirectly involved in managing or accessing healthcare data.

The following stakeholders are chosen:

- **Doctors and Healthcare Professionals/practitioners (DH):** Five doctors are selected from all tier cities and villages who manage patient records through the digital and paper-based modes.
- **Hospital administrative staff (HA):** Five administrative staff are taken from various hospitals from all tier cities and villages, maintaining patient records for future communication.
- **Researchers and Scientist (RS):** Five Researchers from healthcare data analytics and data privacy are selected to work in a new development in healthcare data science and data privacy.
- **Academicians in the computer science field (AC):** Five Academicians in the Computer Science field are selected to either teach data analytics courses or security courses in their curriculum.

The rating obtained from a 1 to 5 Likert scale where one represents strongly agree and 5 represents strongly disagree. Table 11 gives the consolidated ratings about KPIs defined over privacy issues in the Indian healthcare context. It can be seen that almost all the stakeholders are in favor of considering privacy issues.

Table 11. Ratings of KPIs defined over privacy issues by healthcare stakeholders.

KPI	Doctors and Healthcare Professionals/Practitioners					Hospital Administrative Staff					Researchers and Scientist					Academicians in the Computer Science Field					Mean
	DH1	DH2	DH3	DH4	DH5	HA1	HA2	HA3	HA4	HA5	RS1	RS2	RS3	RS4	RS5	AC1	AC2	AC3	AC4	AC5	
Forced Trust Vs. Control	1	2	2	3	3	2	2	2	2	4	1	1	1	2	1	2	2	1	1	2	1.85
Content Viewed by Whom	3	3	4	2	1	3	2	4	1	2	1	1	1	1	1	2	1	3	1	2	1.95
Tacit Knowledge	3	3	3	4	2	2	2	2	3	1	1	1	1	3	2	2	3	2	3	1	2.2
Laws and Regulation	1	1	2	1	2	1	2	1	1	1	1	1	1	1	2	1	3	1	1	3	1.4
Use of new Data Protection Technologies	3	3	3	3	1	3	2	2	3	4	2	2	2	1	1	1	1	1	1	1	2
Researcher's Satisfaction	4	4	4	4	2	5	5	4	4	4	2	2	2	3	1	1	1	1	1	2	2.8
Industry-academia collaboration exists for privacy preservation mechanisms	3	3	3	3	3	3	3	3	3	3	2	2	2	1	1	1	2	1	2	2	2.3

7.2. Future of Data Privacy in India

The government of India took significant steps to address the privacy issues and protect sensitive data from unauthorized access. According to the recent news, the government proposed a law to govern data security in all emerging sectors like healthcare, finance and banking, education, and retail that would give individuals complete ownership of their data. Individuals can access, share, and deny the records associated with them. The Personal Data Protection Bill (PDP) draft was proposed in 2019, which is similar to GDPR [102]. The committee suggested the following key points and developed a privacy framework:

- The law must be flexible and must be adhered to changing technologies.
- Law must be applied to public and private sector entities.
- Entities controlling the data should be accountable for any data processing.
- Consent must be structured and genuine.
- Processing and analysis of data must be minimal.
- Enforcement of the data protection framework should be carried out by a high-powered statutory authority.

Indian published health data are very diverse and collected from different heterogeneous sources, moving towards the healthcare sector and the proposed bill. There are

no regulations over the authorship of the health data, due to which any third party can gain access to the sensitive data and misuse the data. The reidentification attack is the most common attack of health data wherein, with the help of a group of some identifiable entities (called quasi-identifiers), individuals' identities can be easily determined. The proposed law mentioned in the above section has guidelines and technological aspects of preserving healthcare data privacy. The proposed research will be the outcome of the privacy framework developed by the Indian government in the PDP bill. The primary constituents of the PDP bill are drawn in Figure 10.

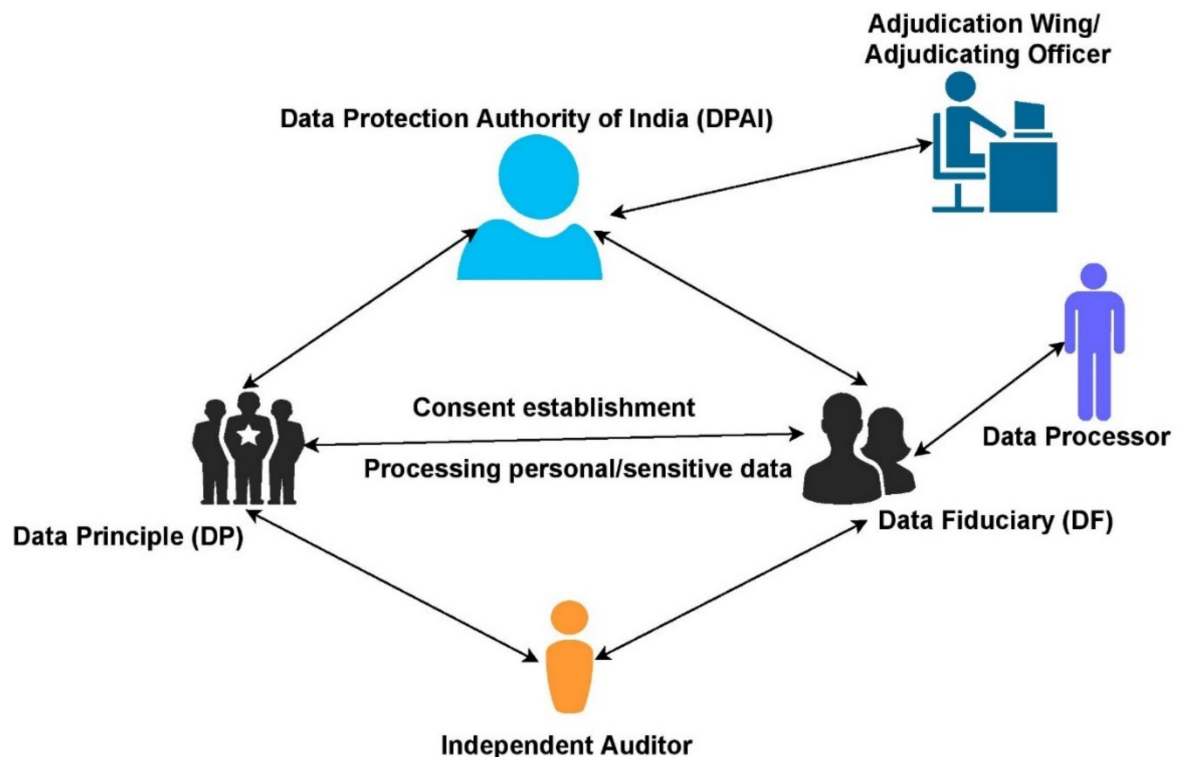


Figure 10. Personal Data Protection bill entities adopted from [101].

Figure 10 presents the essential elements of the data protection bill adopted in India. The data owner is called a data principle in GDPR; they are also called a data custodian. The data fiduciary can be any company, organization, group of people, or individual who determine the purpose of the data use and dissemination. They can be a data holder in the context of GDPR. A data processor is a third-party entity that is involved in the processing of data. In some situations, the data fiduciary and data processor roles can be the same, and the whole depends upon a particular situation. The Data Protection Authority of India (DPAI) is the statutory body that can define rules and regulations about data protection.

7.3. Data Utility-Privacy Trade-Off in Personal Data Protection Bill

Recalling Section 1 of the paper, there is a thin border between data utility and privacy. Privacy is subjective and cannot be fully addressed; data breaches can occur not only due to data publishing but also data pre-processing, data sharing, as well as due to inappropriate policies. Privacy also varies from nation to nation. The newly created PDP bill is enhanced by maintaining the proper balance of data utility and privacy. Recall that Indian healthcare data generated from heterogeneous sources are very unstructured. The user (termed as role) also restricts the access to a particular level of data according to the bill. To provide a different and enhanced level of protection, the data fiduciary can implement data anonymization, randomization, and similar data hiding techniques to ensure that data are protected from various privacy attacks such as background knowledge, linkage,

reidentification, and many more. As per the framework (Figure 10), the data fiduciary and the data processor shall implement the necessary methods such as anonymization and de-identification during data processing to implement appropriate security and safeguard in the system. They also can define what level the privacy can be maintained by keeping the utility of the data. Privacy by Design (PbD) is one of the solutions suggested by DPAI in the PDP bill.

8. Conclusions and Future Work

India is an emerging country in terms of revenue and employment in the healthcare field. The advances of ICT help the healthcare sector streamline data structure and access and health analytics. The healthcare sector in India is growing, although relatively slowly, due to its extensive coverage, strengthening services, and increasing expenditure by public and private investors. This paper covers the overview of the Indian healthcare system with a focus on veracious trends and technologies in healthcare, sectors in healthcare, policies that are driving healthcare systems, and various technologies used in the Indian healthcare system. It is observed that there have been many advances in healthcare in the past year; however, due to unstructured planning, political interventions, and socio-cultural issues, the healthcare system is yet to reach the middle class and poor people in India. The other part of this paper also covered the privacy issues in India. From the existing literature, it is found that major privacy concerns that arise in the Indian healthcare system are due to the doctor–patient relationship, trust management, consent management, lack of security policies, constitutional and political issues, and many more. It can also be deduced that most of the privacy issues in the Indian healthcare system are because of socio-cultural and legal aspects and not because of a lack of technological advancements. In future, more care needs to be taken to improve the attitude of understanding privacy issues for individuals, healthcare professionals, healthcare management, and government rather than improving technological advancements.

Author Contributions: The author P.C. and A.P. wrote the initial draft of the manuscript. A.-J.M.-G. proofread and gave constructive comments towards improving the quality of the paper. He also handled the major revisions in the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This research work is an extensive review of past published papers, and hence no specific data/dataset applies to this research work.

Acknowledgments: The authors would like to thank the anonymous reviewers and editors who have been involved in examining this manuscript.

Conflicts of Interest: The author declares no conflict of interest associated with this research work.

References

1. Healthcare System for New India: Building Blocks—Potential Pathway to Reform. November 2019. Available online: https://niti.gov.in/sites/default/files/2019-11/NitiAayogBook_compressed.pdf (accessed on 1 August 2020).
2. Fullman, N.; Yearwood, J.; Abay, S.M.; Abbafati, C.; Abd-Allah, F.; Abdela, J.; Abdelalim, A.; Abebe, Z.; Abebo, T.A.; Aboyans, V.; et al. Measuring performance on the Healthcare Access and Quality Index for 195 countries and territories and selected subnational locations: A systematic analysis from the Global Burden of Disease Study 2016. *Lancet* **2018**, *391*, 2236–2271. [[CrossRef](#)]
3. Klecun, E. Transforming healthcare: Policy discourses of IT and patient-centred care. *Eur. J. Inf. Syst.* **2016**, *25*, 64–76. [[CrossRef](#)]
4. Top 40 Medical and Healthcare Startups in India. February 2020. Available online: <https://www.medicalstartups.org/country/India/> (accessed on 1 August 2020).
5. Liu, X.; Deng, R.H.; Choo, K.-K.R.; Yang, Y. Privacy-Preserving Reinforcement Learning Design for Patient-Centric Dynamic Treatment Regimes. *IEEE Trans. Emerg. Top. Comput.* **2021**, *9*, 456–470. [[CrossRef](#)]

6. Vora, J.; Devmurari, P.; Tanwar, S.; Tyagi, S.; Kumar, N.; Obaidat, M.S. Blind Signatures Based Secured E-Healthcare System. In Proceedings of the 2018 International Conference on Computer, Information and Telecommunication Systems (CITS), Alsace, Colmar, France, 11–13 July 2018; pp. 1–5.
7. Pussewalage, H.S.G.; Oleshchuk, V.A. Privacy preserving mechanisms for enforcing security and privacy requirements in E-health solutions. *Int. J. Inf. Manag.* **2016**, *36*, 1161–1173. [CrossRef]
8. Dey, N.; Hassanién, A.E.; Bhatt, C.; Ashour, A.; Satapathy, S.C. (Eds.) *Internet of Things and Big Data Analytics toward Next-Generation Intelligence*; Springer: Berlin, Germany, 2018; pp. 3–549.
9. Kaur, G.; Tomar, P.; Singh, P. Design of Cloud-Based Green IoT Architecture for Smart Cities. In *Internet of Things and Big Data Analytics toward Next-Generation Intelligence*; Springer: Cham, Switzerland, 2018; pp. 315–333.
10. Valdez, A.C.; Ziefle, M. The users' perspective on the privacy-utility trade-offs in health recommender systems. *Int. J. Hum. Comput. Stud.* **2019**, *121*, 108–121. [CrossRef]
11. Churi, P.P.; Pawar, A.V. *J. Eng. Sci. Technol. Rev.* **2019**, *12*, 17–25. [CrossRef]
12. Sánchez, D.; Batet, M.; Viejo, A. Utility-preserving privacy protection of textual healthcare documents. *J. Biomed. Inform.* **2014**, *52*, 189–198. [CrossRef]
13. Rastogi, V.; Suciú, D.; Hong, S. The boundary between privacy and utility in data publishing. In Proceedings of the 33rd International Conference on Very Large Data Bases, Vienna, Austria, 23–28 September 2007; pp. 531–542.
14. Talib, F.; Rahman, Z. Current Health of Indian Healthcare and Hospitality Industries: A Demographic Study. *Int. J. Bus. Res. Dev.* **2013**, *2*, 1–17. [CrossRef]
15. Available online: <https://www.ibef.org/industry/healthcare-india.aspx> (accessed on 13 June 2020).
16. Sachdeva, S.; Batra, S.; Bhalla, S. Evolving large scale healthcare applications using open standards. *Health Policy Technol.* **2017**, *6*, 410–425. [CrossRef]
17. Available online: [http://www.cbhidghs.nic.in/Ebook/National%20Health%20Profile-2018%20\(e-Book\)/files/assets/common/downloads/files/NHP%202018.pdf](http://www.cbhidghs.nic.in/Ebook/National%20Health%20Profile-2018%20(e-Book)/files/assets/common/downloads/files/NHP%202018.pdf) (accessed on 14 June 2020).
18. Ganesan, L.; Veena, S.R. 'Make In India' For Healthcare Sector in India: A SWOT Analysis on Current Status and Future Prospects. *Int. J. Health Sci. Res.* **2018**, *8*, 258–265.
19. Shankar, D. Health sector reforms for 21st century healthcare. *J. Ayurveda Integr. Med.* **2015**, *6*, 4. [CrossRef]
20. IRDA, CII, Grant Thornton, Gartner, Technopak, TechSci Research. Available online: <https://www.ibef.org/download/Healthcare-July-2019.pdf> (accessed on 14 June 2020).
21. Available online: https://mea.gov.in/images/pdf/22899_India_in_Business.pdf (accessed on 14 June 2020).
22. Basu, S.; Andrews, J.; Kishore, S.; Panjabi, R.; Stuckler, D. Comparative performance of private and public healthcare systems in low-and middle-income countries: A systematic review. *PLoS Med.* **2012**, *9*, e1001244. [CrossRef]
23. Gupta, Y.; Joshi, A.; Kale, G. Healthcare Analytics Systems: An Overview. *Int. J. Eng. Sci.* **2018**, *8*, 18898.
24. Boric-Lubecke, O.; Gao, X.; Yavari, E.; Baboli, M.; Singh, A.; Lubecke, V.M. E-healthcare: Remote monitoring, privacy, and security. In Proceedings of the 2014 IEEE MTT-S International Microwave Symposium (IMS2014), Tampa, FL, USA, 1–6 June 2014; pp. 1–3.
25. Morilla, M.D.R.; Sans, M.; Casasa, A.; Giménez, N. Implementing technology in healthcare: Insights from physicians. *BMC Med. Inform. Decis. Mak.* **2017**, *17*, 92. [CrossRef] [PubMed]
26. Kapoor, V.; Singh, R.; Reddy, R.; Churi, P. Privacy Issues in Wearable Technology: An Intrinsic Review. 2020. SSRN 3566918. Available online: <https://ssrn.com/abstract=3566918> (accessed on 20 June 2021). [CrossRef]
27. Vassy, J.L.; Korf, B.R.; Green, R.C. How to know when physicians are ready for genomic medicine. *Sci. Transl. Med.* **2015**, *7*, 287fs19. [CrossRef] [PubMed]
28. Chen, X.; Ishwaran, H. Random forests for genomic data analysis. *Genomics* **2012**, *99*, 323–329. [CrossRef]
29. Das, A.K.; Kedia, A.; Sinha, L.; Goswami, S.; Chakrabarti, T.; Chakrabarti, A. Data mining techniques in Indian healthcare: A short review. In Proceedings of the 2015 International Conference on Man and Machine Interfacing (MAMI), Bhubaneswar, India, 17–19 December 2015; pp. 1–7.
30. Pandey, A.; Roy, N.; Bhawsar, R.; Mishra, R.M. Health Information System in India: Issues of data availability and quality. *Demogr. India* **2010**, *39*, 111–128.
31. Patel, S.; Patel, H. Survey of Data Mining Techniques used in Healthcare Domain. *Int. J. Inf.* **2016**, *6*, 53–60. [CrossRef]
32. Rojas, E.; Munoz-Gama, J.; Sepúlveda, M.; Capurro, D. Process mining in healthcare: A literature review. *J. Biomed. Inform.* **2016**, *61*, 224–236. [CrossRef]
33. Tomar, D.; Agarwal, S. A survey on Data Mining approaches for Healthcare. *Int. J. Bio-Sci. Bio-Technol.* **2013**, *5*, 241–266. [CrossRef]
34. Kalaiselvi, C. Diagnosing of heart diseases using average k-nearest neighbor algorithm of data mining. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–18 March 2016; pp. 3099–3103.
35. Kala, V.C.; Kiran, L.V.; Prasad, P.S. Prediction of diseases with pathological characteristics classification using data mining. In Proceedings of the 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), Vellore, India, 30–31 March 2019; pp. 1–5.
36. Chauhan, D.; Jaiswal, V. An efficient data mining classification approach for detecting lung cancer disease. In Proceedings of the 2016 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 21–22 October 2016; pp. 1–8.

37. Khan, S.I.; Islam, A.; Hossen, A.; Zahangir, T.I.; Hoque, A.S.M.L. Supporting the Treatment of Mental Diseases using Data Mining. In Proceedings of the 2018 International Conference on Innovations in Science, Engineering and Technology (ICISSET), Chittagong, Bangladesh, 27–28 October 2018; pp. 339–344.
38. Alonso, S.G.; De La Torre-Díez, I.; Hamrioui, S.; López-Coronado, M.; Barreno, D.C.; Nozaleda, L.M.; Franco, M. Data Mining Algorithms and Techniques in Mental Health: A Systematic Review. *J. Med. Syst.* **2018**, *42*, 161. [[CrossRef](#)]
39. Dubey, A.K.; Gupta, U.; Jain, S. Epidemiology of lung cancer and approaches for its prediction: A systematic review and analysis. *Chin. J. Cancer* **2016**, *35*, 71. [[CrossRef](#)]
40. Swapna, K.; Babu, M.S.P. A Critical Study on Cluster Analysis Methods to Extract Liver Disease Patterns in Indian Liver Patient Data. *Int. J. Comput. Intell. Res.* **2017**, *13*, 2379–2390.
41. Pasha, M.; Fatima, M. Comparative Analysis of Meta Learning Algorithms for Liver Disease Detection. *J. Softw.* **2017**, *12*, 923–933. [[CrossRef](#)]
42. Abdar, M.; Zomorodi-Moghadam, M.; Das, R.; Ting, I.-H. Performance analysis of classification algorithms on early detection of liver disease. *Expert Syst. Appl.* **2017**, *67*, 239–251. [[CrossRef](#)]
43. Wu, H.; Yang, S.; Huang, Z.; He, J.; Wang, X. Type 2 diabetes mellitus prediction model based on data mining. *Inform. Med. Unlocked* **2018**, *10*, 100–107. [[CrossRef](#)]
44. Verma, L.; Srivastava, S.; Negi, P.C. A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data. *J. Med. Syst.* **2016**, *40*, 178. [[CrossRef](#)] [[PubMed](#)]
45. Goswami, T.; Dabhi, V.K.; Prajapati, H.B. Skin Disease Classification from Image—A Survey. In Proceedings of the 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 6–7 March 2020; pp. 599–605.
46. Induja, S.; Raji, C. Computational Methods for Predicting Chronic Disease in Healthcare Communities. In Proceedings of the 2019 International Conference on Data Science and Communication (IconDSC), Bangalore, India, 1–2 March 2019; pp. 1–6.
47. Jain, D.; Singh, V. Feature selection and classification systems for chronic disease prediction: A review. *Egypt. Inform. J.* **2018**, *19*, 179–189. [[CrossRef](#)]
48. Chaurasia, V.; Pal, S. A novel approach for breast cancer detection using data mining techniques. *Int. J. Innov. Res. Comput. Commun. Eng.* **2017**, *2*.
49. Kadam, K.; Kamat, P.V.; Malav, A.P. Cardiovascular Disease Prediction Using Data Mining Techniques: A Proposed Framework Using Big Data Approach. In *Coronary and Cardiothoracic Critical Care: Breakthroughs in Research and Practice*; IGI Global: Hershey, PA, USA, 2019; pp. 246–264.
50. Miljkovic, D.; Aleksovski, D.; Podpečan, V.; Lavrač, N.; Malle, B.; Holzinger, A. Machine Learning and Data Mining Methods for Managing Parkinson’s disease. In *Machine Learning for Health Informatics*; Springer: Cham, Switzerland, 2016; pp. 209–220.
51. Manogaran, G.; Lopez, D. A survey of big data architectures and machine learning algorithms in Healthcare. *Int. J. Biomed. Eng. Technol.* **2017**, *25*, 182–211. [[CrossRef](#)]
52. Yu, K.-H.; Beam, A.L.; Kohane, I.S. Artificial intelligence in healthcare. *Nat. Biomed. Eng.* **2018**, *2*, 719–731. [[CrossRef](#)] [[PubMed](#)]
53. Simonite, T. Google’s AI Eye Doctor Gets Ready to Go to Work in India. WIRED. Available online: <https://www.wired.com/2017/06/googles-ai-eye-doctor-gets-ready-go-work-india/> (accessed on 18 June 2020).
54. Waghade, S.S.; Karandikar, A.M. A comprehensive study of healthcare fraud detection based on machine learning. *Int. J. Appl. Eng. Res.* **2018**, *13*, 4175–4178.
55. Baid, U.; Baheti, B.; Dutande, P.; Talbar, S. Detection of Pathological Myopia and Optic Disc Segmentation with Deep Convolutional Neural Networks. In Proceedings of the TENCON 2019-2019 IEEE Region 10 Conference (TENCON), Kochi, India, 17–20 October 2019; pp. 1345–1350.
56. Mandal, T.; Rao, K.S. Glottal Closure Instants Detection From Pathological Acoustic Speech Signal Using Deep Learning. *arXiv* **2018**, arXiv:1811.09956.
57. Gupta, N.; Ahuja, N.; Malhotra, S.; Bala, A.; Kaur, G. Intelligent heart disease prediction in cloud environment through ensembling. *Expert Syst.* **2017**, *34*, e12207. [[CrossRef](#)]
58. Abdar, M.; Yen, N.Y.; Hung, J.C.-S. Improving the Diagnosis of Liver Disease Using Multilayer Perceptron Neural Network and Boosted Decision Trees. *J. Med. Biol. Eng.* **2018**, *38*, 953–965. [[CrossRef](#)]
59. Patnaik, S.K.; Sidhu, M.S.; Gehlot, Y.; Sharma, B.; Muthu, P. Automated Skin Disease Identification using Deep Learning Algorithm. *Biomed. Pharmacol. J.* **2018**, *11*, 1429–1437. [[CrossRef](#)]
60. Kaur, H.; Kumari, V. Predictive modelling and analytics for diabetes using a machine learning approach. *Appl. Comput. Inform.* **2020**. [[CrossRef](#)]
61. Kumar, R.; Arora, R.; Bansal, V.; Sahayasheela, V.J.; Buckchash, H.; Imran, J.; Narayanan, N.; Pandian, G.N.; Raman, B. Accurate Prediction of COVID-19 using Chest X-Ray Images through Deep Feature Learning model with SMOTE and Machine Learning Classifiers. *medRxiv* **2020**. [[CrossRef](#)]
62. Kakileti, S.T.; Madhu, H.J.; Manjunath, G.; Wee, L.; Dekker, A.; Sampangi, S. Personalized risk prediction for breast cancer pre-screening using artificial intelligence and thermal radiomics. *Artif. Intell. Med.* **2020**, *105*, 101854. [[CrossRef](#)]
63. Ko, I.; Chang, H. Interactive data visualization based on conventional statistical findings for antihypertensive prescriptions using National Health Insurance claims data. *Int. J. Med. Inform.* **2018**, *116*, 1–8. [[CrossRef](#)]

64. Ledesma, A.; Bidargaddi, N.; Strobel, J.; Schrader, G.; Nieminen, H.; Korhonen, I.; Ermes, M. Health timeline: An insight-based study of a timeline visualization of clinical data. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 170. [[CrossRef](#)]
65. Ulahannan, J.P.; Narayanan, N.; Thalath, N.; Prabhakaran, P.; Chaliyeduth, S.; Suresh, S.P.; Mohammed, M.; Rajeevan, E.; Joseph, S.; Balakrishnan, A.; et al. A citizen science initiative for open data and visualization of COVID-19 outbreak in Kerala, India. *J. Am. Med. Inform. Assoc.* **2020**, *27*, 1913–1920. [[CrossRef](#)]
66. Tandon, H.; Ranjan, P.; Chakraborty, T.; Suhag, V. Coronavirus (COVID-19): ARIMA based time-series analysis to forecast near future. *arXiv* **2020**, arXiv:2004.07859.
67. Gupta, S.; Raghuwanshi, G.S.; Chanda, A. Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020. *Sci. Total Environ.* **2020**, *728*, 138860. [[CrossRef](#)]
68. Basu, S.; Mitra, S.; Saha, N. Deep Learning for Screening COVID-19 using Chest X-Ray Images. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, ACT, Australia, 1–4 December 2020; pp. 2521–2527.
69. Pasa, F.; Golkov, V.; Pfeiffer, F.; Cremers, D. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization. *Sci. Rep.* **2019**, *9*, 1–9. [[CrossRef](#)]
70. Olshannikova, E.; Ometov, A.; Koucheryavy, Y.; Olsson, T. Visualizing Big Data with augmented and virtual reality: Challenges and research agenda. *J. Big Data* **2015**, *2*, 22. [[CrossRef](#)]
71. Schuemie, M.J.; Van Der Straaten, P.; Krijn, M.; Van Der Mast, C.A. Research on Presence in Virtual Reality: A Survey. *Cyberpsychol. Behav.* **2001**, *4*, 183–201. [[CrossRef](#)] [[PubMed](#)]
72. Prakaash, D.; Kodagahally, R.R.; Honnaiah, M. Virtual reality: A railroad for structural bioinformatics towards advanced cancer research. *PeerJ Prepr.* **2017**, *5*, e2960v1.
73. Ayoub, A.; Puljiala, Y. The application of virtual reality and augmented reality in Oral & Maxillofacial Surgery. *BMC Oral Health* **2019**, *19*, 238. [[CrossRef](#)]
74. Chamberlain, D.; Jimenez-Galindo, A.; Fletcher, R.R.; Kodgule, R. Applying Augmented Reality to Enable Automated and Low-Cost Data Capture from Medical Devices. In Proceedings of the Eighth International Conference on Information and Communication Technologies and Development; ACM: New York, NY, USA, 2016; pp. 1–4.
75. Bhatti, R.; Grandison, T. Towards Improved Privacy Policy Coverage in Healthcare Using Policy Refinement. In *Workshop on Secure Data Management*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 158–173.
76. Parker, R.B. A definition of privacy. *Rutgers L. Rev.* **1973**, *27*, 275.
77. Mireku, K.K.; Zhang, F.; Komlan, G.; Kingsford, K.M.; FengLi, Z. Patient knowledge and data privacy in healthcare records system. In Proceedings of the 2017 2nd International Conference on Communication Systems, Computing and IT Applications (CSCITA), Mumbai, India, 7–April 2017; pp. 154–159.
78. Cheng, V.S.; Hung, P.C. Health Insurance Portability and Accountability Act (HIPPA) Compliant access control model for web services. *Int. J. Healthc. Inf. Syst. Inform.* **2006**, *1*, 22–39. [[CrossRef](#)]
79. Cho, S. A Study on Privacy Protection in the EU’s GDPR and Korea’s Personal Information Protection Act. *Law J.* **2018**, *61*, 117–148.
80. Akpojivi, U. Rethinking information privacy in a “connected” world. In *Censorship, Surveillance, and Privacy: Concepts, Methodologies, Tools, and Applications*; IGI Global: Hershey, PA, USA, 2019; pp. 1–18.
81. Churi, P.P.; Pawar, A.V. A Systematic Review on Privacy Preserving Data Publishing Techniques. *J. Eng. Sci. Technol. Rev.* **2019**, *12*, 17–25. [[CrossRef](#)]
82. Srinivas, N.; Biswas, A. Protecting patient information in India: Data privacy law and its challenges. *NUJS L. Rev.* **2012**, *5*, 411.
83. Jalali, M.S.; Landman, A.; Gordon, W.J. Telemedicine, privacy, and information security in the age of COVID-19. *J. Am. Med. Inform. Assoc.* **2021**, *28*, 671–672. [[CrossRef](#)]
84. Kagalwalla, N.; Garg, T.; Churi, P.; Pawar, A. A survey on implementing Privacy in Healthcare: An indian perspective. *Int. J. Adv. Trends Comput. Sci. Eng.* **2019**, *8*, 963–982. [[CrossRef](#)]
85. Meingast, M.; Roosta, T.; Sastry, S. Security and privacy issues with health care information technology. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 5453–5458.
86. Mishra, N.N.; Parker, L.S.; Nimgaonkar, V.L.; Deshpande, S.N. Privacy and the Right to Information Act, 2005. *Indian J. Med. Ethics* **2008**, *5*, 158–161. [[CrossRef](#)] [[PubMed](#)]
87. Berger, D. Corruption ruins the doctor-patient relationship in India. *BMJ* **2014**, *348*, g3169. [[CrossRef](#)] [[PubMed](#)]
88. Ion, I.; Sachdeva, N.; Kumaraguru, P.; Čapkun, S. Home is safer than the cloud! Privacy concerns for consumer cloud storage. In Proceedings of the Seventh Symposium on Usable Privacy and Security; 2011; pp. 1–20. [[CrossRef](#)]
89. Priya, R.; Sivasankaran, S.; Ravisasthiri, P.; Sivachandiran, S. A survey on security attacks in electronic healthcare systems. In Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCCSP), Chennai, India, 6–8 April 2017; pp. 0691–0694.
90. Kumar, P.R.; Raj, P.H.; Jelciana, P. Exploring Data Security Issues and Solutions in Cloud Computing. *Procedia Comput. Sci.* **2018**, *125*, 691–697. [[CrossRef](#)]
91. Carol, Y.; Cheung, Y. Data without borders. *Lancet* **2019**, *393*, 1331–1384.
92. Bartoletti, I. AI in Healthcare: Ethical and Privacy Challenges. In *Conference on Artificial Intelligence in Medicine in Europe*; Springer: Cham, Switzerland, 2019; pp. 7–10.

93. Bhattacharya, P.; Tanwar, S.; Bodke, U.; Tyagi, S.; Kumar, N. BinDaaS: Blockchain-Based Deep-Learning as-a-Service in Healthcare 4.0 Applications. *IEEE Trans. Netw. Sci. Eng.* **2019**, *1*. [[CrossRef](#)]
94. Hathaliya, J.J.; Tanwar, S.; Tyagi, S.; Kumar, N. Securing electronics healthcare records in Healthcare 4.0: A biometric-based approach. *Comput. Electr. Eng.* **2019**, *76*, 398–410. [[CrossRef](#)]
95. Kumari, A.; Tanwar, S.; Tyagi, S.; Kumar, N. Fog computing for Healthcare 4.0 environment: Opportunities and challenges. *Comput. Electr. Eng.* **2018**, *72*, 1–13. [[CrossRef](#)]
96. Vora, J.; Nayyar, A.; Tanwar, S.; Tyagi, S.; Kumar, N.; Obaidat, M.S.; Rodrigues, J.J.P.C. BHEEM: A Blockchain-Based Framework for Securing Electronic Health Records. In Proceedings of the 2018 IEEE Globecom Workshops (GC Wkshps), Abu Dhabi, United Arab Emirates, 9–13 December 2018; pp. 1–6.
97. Vora, J.; Tanwar, S.; Tyagi, S.; Kumar, N.; Rodrigues, J.J.P.C. FAAL: Fog computing-based patient monitoring system for ambient assisted living. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017; pp. 1–6.
98. Vora, J.; Tanwar, S.; Tyagi, S.; Kumar, N.; Rodrigues, J.J.P.C. Home-based exercise system for patients using IoT enabled smart speaker. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017; pp. 1–6.
99. Hathaliya, J.J.; Tanwar, S. An exhaustive survey on security and privacy issues in Healthcare 4. *Comput. Commun.* **2020**, *153*, 311–335. [[CrossRef](#)]
100. Hathaliya, J.J.; Tanwar, S.; Evans, R. Securing electronic healthcare records: A mobile-based biometric authentication approach. *J. Inf. Secur. Appl.* **2020**, *53*, 102528. [[CrossRef](#)]
101. Madhisetty, S.; Williams, M.-A. Managing Privacy Through Key Performance Indicators When Photos and Videos Are Shared via Social Media. In *Science and Information Conference*; Springer: Cham, Switzerland, 2018; pp. 1103–1117.
102. Singh, R.G.; Ruj, S. A Technical Look at the Indian Personal Data Protection Bill. *arXiv* **2020**, arXiv:2005.13812.