Visualising Developing Nations Health Records: Opportunities, Challenges and Research Agenda

Afamefuna Umejiaku  
Texas Tech University,  
United States  
anumejiaku@gmail.com

Tommy Dang  
Texas Tech University,  
United States  
tommy.dang@ttu.edu

ABSTRACT

The benefits of effectively visualizing health records in huge volumes has resulted in health organizations, insurance companies, policy and decision makers, governments and drug manufacturers’ transformation in the way research is conducted. This has also played a key role in determining investment of resources. Health records contain highly valuable information; processing these records in large volumes is now possible due to technological advancement which allows for the extraction of highly valuable knowledge that has resulted in breakthroughs in scientific communities. To visualize health records in large volumes, the records need to be stored in electronic forms, properly documented, processed, and analyzed. A good visualization technique is used to present the analyzed information, allowing for effective knowledge extraction which is done in a secured manner protecting the privacy of the patients whose health records were used. As research and technological advancement have improved, the quality of knowledge extracted from health records have also improved; unfortunately, the numerous benefits of visualizing health records have only been felt in developed nations, unlike other sectors where technological advancement in developed nations have had similar impact in developing nations. This paper identifies the characteristics of health records and the challenges involved in processing large volumes of health records. This is to identify possible steps that could be taken for developing nations to benefit from visualizing health records in huge volumes.

KEYWORDS

Visualisation, Health records, Developing Nations

ACM Reference Format:
Afamefuna Umejiaku and Tommy Dang. 2021. Visualising Developing Nations Health Records: Opportunities, Challenges and Research Agenda. In The 12th International Conference on Advances in Information Technology (IAIT2021), June 29-July 1, 2021, Bangkok, Thailand. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3468784.3471607

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

IAIT2021, June 29-July 1, 2021, Bangkok, Thailand
© 2021 Association for Computing Machinery.
ACM ISBN 978-1-4503-9012-5/21/06… $15.00
https://doi.org/10.1145/3468784.3471607

1 INTRODUCTION

The global pandemic caused by the Coronavirus in 2020 can be seen as an indication that more needs to be done to support the development of healthcare platforms that transcend national or state borders. Health records are private, yet contain information that could be very beneficial to the patient and society at large when analyzed [7, 47, 59]. Therapeutics, monitoring, treatment and care of patients in the healthcare industry produce vast amount of data currently estimated to be in Exabytes yearly, [47] applying techniques in big data research to health care data has been shown to improve understanding of health challenges, its monitoring and prediction [47, 60, 61]. Big data research has received a lot of attention in recent years [40] and it been listed in top 10 Critical and strategic technology trends in the last decade [39]. Kalantari et al [27] pointed out in their bibliometric analysis of big data that top countries are publishing a vast number of articles in the field while there is an essential lack of interest in the field in 96 other countries with no publication. Big data research has proven to be very beneficial, transforming the way several problems are currently being approached in several fields including healthcare [35], a key area in big data research is its visualisation. Big data visualisation is the graphical representation of data using visual elements providing means for the extraction of information and enabling one to make broad predictions [28, 35].

Visualising health record provides useful insight with numerous benefits to the society [4, 19]. Visualization as an effective tool in disseminating big data has been gaining traction in the health sector because it protects the identity of patients as the data isn’t viewed individually but collectively and it allows both scientist and the general audience to understand health data, it also detects hidden patterns and correlations, [8, 54] and spots meanings that other machine techniques might not spot. By visualizing health records of a locality, predictions of a disease outbreak can be determined faster; drug behaviors and any potential outbreak of an infection could also be quickly noticed and effectively monitored on a large scale.

This survey paper identifies the opportunities and challenges of visualizing health records in developing nations. Like never before, people are not only bothered about their own health, loved ones’ health or their own nations’ health outlook but they are also concerned about the health situation in other nations, raising concerns about the accuracy of what is being reported in developing nations [16]. Just as big data changed the way we do business and management in large organizations [39], having national and global visualization platforms in the healthcare sector could transform the way healthcare research is performed, the way funding is channelled, and the way future pandemics could be prevented. By seeking to
understand and identify challenges that could be unique to developing nations and identifying factors hindering the development of a reliable large scale or centralized visualization strategies, solutions could be proffered to address these challenges. Our contributions in this study are:

(1) We broadly identified the challenges involved in Visualizing patients’ health records and steps that have been taken so far to address these challenges.

(2) We identified challenges unique to developing nations with regards to the handling of health records and the dissemination of information from health records using visualization techniques in the healthcare sector.

Big data and electronic health record have received lots of attention over the past decade but little has been done in developing nations [27]. This paper therefore focuses on ensuring that developing nations, irrespective of the challenges they face locally, could utilize some of the advancements made in the technological field to the health care sector enhancing the use of visualization to extract useful knowledge and ensure the dissemination of information. This paper is organized as follows: Section 1 is the introduction, section 2 provides an overview of health records, how they relate to big Data and visualization of health records. Health records characteristics are discussed in section 3, while section 4 focuses on processing health records. Management challenges are discussed in section 5; Section 6 discusses challenges that are unique to developing nations and section 7 is the conclusion.

2 OVERVIEW

Governments and individuals spend a lot of resources in the health sector, a challenging service affecting patients’ life and economies. [25] Information is extracted from patients to help improve their health situation and this information can also be used for various other services; as such the manner in which the information is documented and stored is important. The use of electronic technologies to improve health care has been embraced over the years due to its numerous benefits, and health providers have been encouraged to adopt electronic health records though several health providers especially in developing nations, are yet to do so. Extracting valuable information form these records could have significant medical and economic benefits. [5, 6].

2.1 Health Records

Health records are records generated from patients diagnostic services and stored by health professionals or the use of sensors attached to persons. They contain a person’s social demographics, family history, lab reports, imaging data, genomics data, endoscopic data, colonoscopies data, clinical notes, etc. and are estimated to currently be in Zettabyte [47]. These records are not just big, they are complex, in different formats and could be structured or unstructured.

Health record can be divided [3] into two parts - the creation part and the access part. Creation part involves the direct interaction of patients with healthcare professionals or via the use of electronic data capture devices such as sensors. It also involves how the data is formatted as well as the storage medium. Various standards and policies have been established to ensure the records created meet required expectations. Access part is responsible for who has access to the data stored, ensuring the security and confidentiality of health records.

Challenges associated with health records [47] include -

- Correctness - How accurate is the collected data often a challenge, as incorrect data could lead to wrong outcomes.
- Incompleteness - Patients health records are often incomplete.
- Inconsistency - Multiple person associated with the creation of health records often lead to inconsistency.
- Privacy - Privacy concerns are often raise as health records may be shared or linked without consent or authorisation.
- security - The vulnerability of health data is often an issue as criminals often target these data as the contain highly valuable information.

Diverse numbers of systems have been developed to handle the management and storage of these health records once created. Hospitals, laboratories, Radiology, Clinics among others have all developed their systems to handle such records [59]. The International Organization for Standardization (ISO) Committee Draft Technical Recommendation 20514, ISO defined Electronic Health Records (EHR) broadly as “a repository of information regarding the health status of a subject of care in a computer process-able form, stored and transmitted securely and accessible by multiple authorized users”. Making the most from health data would involve several crucial but important steps that every health provider would need to partake in as described in section 6.

2.2 Health records as big data

Big data definition as proposed by Andrea et al “is the Information asset characterized by such a high volume, velocity and variety to require specific technology and analytical methods for its transformation into value” [15], Radhika et al [40] suggested that a 100MB file to be sent via email could be considered big data since attachment capacity for email is restricted to 25MB. Seagate’s report predicts that by 2025 data creation would swell to 163 zettabytes(ZB) [14], Electronic health records have witnessed tremendous growth, from 500 petabytes in 2012 to several zettabytes(ZB)in 2020 [35, 59]. With the massive amount of data currently being captured from health records, the current challenge is extracting meaningful information that could be acted upon. Uncertainty however remains about the use of big data analytics on health records especially with regards to its governance. [6, 60]

2.3 Visualizing Health records

Health records visualization entails the use of tools that produce graphics, images, diagrams or animation for capturing information from health records, displaying an easily understandable information and allowing for interaction with the data thus facilitating the discovery of knowledge [35, 59]. A good health records visualization allows for quick analysis of health records and visual story-telling capturing the attention of an audience and allowing for easy extraction of knowledge at a faster rate with less mental work. [13, 35] A good health records visualization technique requires the use of an effective and interesting visualization type.
Examples of visualization techniques are word cloud, dashboards, charts, tables, graphs, maps, info-graphics, area chart, bar chart, box-and-whisker plots, bubble cloud, bullet graph, circle view, dot distribution map, heat map, histogram, matrix, polar area, radial tree, scatter plot (2D or 3D) and tree-map.

To create a Good health records Visualization [33, 52, 53], the designer must consider in the following:

- what are the questions they targeted individuals want to solve.
- what questions should the visualization answer.
- what other question could be inspired by visualization.
- how do I get the targeted individuals inspired.

Once these questions have been answered, the designer must then

- Choose the right visualization technique for the job. For example, histograms conveniently show clustered data while bar charts are effective for comparison.
- Make it meaningful, choosing a pattern that could be easily followed, passing as much information as possible at a glance [13].
- Incorporate clues using colour, shapes, designs, size and text carefully and intentionally.
- Eliminate clutter.
- Add interactivity ensuring it allows for qualitative and quantitative analysis
- Ensure the latency is as low as possible as delays in displaying a visualisation could result in loss of interest.

To facilitate analysis of big data like health records several techniques have been developed [19]; these include FlsViz, WiFlsViz, FpVAT, PyramidViz, FpMapViz and HSLviz. The use of Virtual reality [54] has also been explored and some of the most popular visualization tools [34, 44] are Tableau Software, Qlik and TIBCO Software, Zoho Analytics, IBM Cognos Analytics, Microsoft Power BI, Plotly, Gephi and Oracle Visual Analyzer.

Benefit of Visualizing health records include

- **Risk factor identification**: Identifying possible risk is often a challenge; visualization however makes it easy to observe risk trends.
- **Survival prediction and mortality**: By simply observing trends in health data visualization, prediction of risk without much calculations is easier.
- **readmission or length of stay prediction**: Hospital readmission is common and expensive [18], predicting potentially avoidable readmission benefits both the patient and aids in the effective utilization of healthcare resources.
- **Early warning**: Visualization tools are very powerful tools for displaying the cause–effect relationships between different events allowing for anticipation of future occurrences [11].
- **Knowledge discovery**: Visualization often exposes trends that would not have been observed by just applying mathematical solution. This often leads to new discoveries.
- **Study of drug behaviors with different diseases or different patients**: The impact of a new drug is not always obvious as one might not easily see all the possible implications of the drug use. Visualization however gives us a quick mental picture of the likely impact of a drug. The impact of a new drug is not always obvious as one might not easily see all the possible implications of the drug use. Visualization however gives us a quick mental picture of the likely impact of a drug.

### 3 HEALTH RECORDS CHARACTERISTICS

Health records characteristics arise from the life cycle of the health record. Just as different researchers have distinct understanding on the characteristics of big data, health records characteristics can be described using 3V’s, 5V’s and 7V’s with each V representing a unique characteristics of the health record. The V’s represent a characteristics of the record where 3V’s often describes volume, velocity and variety, 5V’s describe velocity, volume, variety, veracity and variability. The 7V’s are discussed below [50]:

#### 3.1 Volume

Volume is the total amount of health records produced measured in bytes. It is impractical to define a specific threshold because as technology improves so does storage capacity increase allowing for the storage of even more data. Processing speeds have also been increasing allowing for the processing of larger amount of data in real-time; as a result, what we may is consider big data today may not be considered big data in the future [21].

#### 3.2 Velocity

Velocity is the rate at which health records is generated, processed, and analyzed. Analyzing unstructured data on a large scale was a challenge for most organizations hence they were just stored in their databases and analyzed in bits but with technological advancements, especially machine learning techniques, more valuable information are being extracted from both structured and unstructured data at
a much faster rate. The use of sensors and other electronic data capture devices have increased the amount of data being generated hence real-time analysis of health record is a tedious task especially for time sensitive data [21].

3.3 Variety
Variety deals with the structural heterogeneity’s of data, the fact that health records are generated and stored in various formats that may be structured, semi-structured or unstructured. Structured health records include laboratory results, vital signs, International Classification of Diseases (ICD) code. Unstructured data include narrative information (free text) such as images and graphics, visit notes, discharge summary, chief complaint. Semi-structured materials include prescriptions and medications [20].

3.4 Variability
Variability deals with different means health records could possess even if the look similar. Trying to analyze health records in bits could lead to wrong conclusions. However, combination of several forms of information provides for better understanding of the situation. For example, a doctors’ report for a patient and a few laboratory results could inform a decision on a likely treatment course but once the patient provides information about their family history, a different treatment course is explored. Thus, accounting for variability in health records is very important [50].

3.5 Veracity
Veracity is the accuracy of health records. It entails more than just the data quality but an understanding of the data, discovering discrepancies as health records often are incomplete and contain inaccurate information [47, 50].

3.6 Visualisation
Visualising health records enables the capturing of meaningful information from health records as has been described in Section 2.3.

3.7 Value
Value includes actionable information, knowledge or any useful information obtained from health records. It is basically the reason for health records. Health records are very valuable as they are a repository of information that once extracted often have significant benefits to the society. Other than its primary use in patient personal care, knowledge that is extracted could be beneficial for planning future services, resource allocation, developing standards, preventive campaigns, health facilities audit, clinical research, amongst others. [47].

4 HEALTH RECORDS PROCESSING CHALLENGES
This section explores the challenges involved in processing health records to obtain valuable information. Health records Processing begins with the capture of patients’ information, storing the information, processing the stored information and finally extracting knowledge from the information. Researchers have proposed solutions to several of the challenges in processing health records as seen in Fig 1. These challenges are further discussed below and are often similar to those encountered in processing big data. [50].

![Figure 1: Health records processing](image-url)

4.1 Health Records Generation and Storage
Health records are generated from several sources, primarily these records are generated at the instance of a health professional in a health facility following some set standards or guidelines, how health records are generated and stored determine the ease of extracting information from them. Some health organizations use electronic means to store information, and these often have numerous benefits as opined by Shahid & Rizwan. Electronic devices tend to always follow a similar pattern or guideline as directed by the device being used and part of the information extracted may still be unstructured, there is still an easy way to follow structure hence processing data generated and stored electronically is faster more accurate and widely adopted in developed nations. However, several developing nations still store health records manually often using paper. Paper based records face problems such as records getting missing, alterations and difficulty in transporting or sharing [26, 45]. Patil, Sunil et al [51] conducted a survey to recognize the benefits of using electronic devices to store health records and about three-quarters of those surveyed opined that using electronic devices is important. The early 1990’s witnessed the introduction of electronic health records and though there have been enormous advancements in the field, Evans R [17] opined that a lot still needs to be done as expectation have not been met and the needs of today’s rapidly changing world, have not been met.
Visualising Developing Nations Health Records: Opportunities, Challenges and Research Agenda IAIT2021, June 29-July 1, 2021, Bangkok, Thailand

4.2 Health records cleaning, Integration, Aggregation and mining

Cleaning is a process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a data-set. Integration entails the merging of health records into a centralized system or creation of approaches that support the use of distributed systems facilitating access to information by authorized individuals. Aggregation deals with preparing integrated health records for information processing and data. Mining is the process of extracting knowledge or actionable information from raw data, all these four elements are essential in the preparation of health records [50].

Health records, irrespective of the storage platform being used, are often heterogeneous: they may be structured, unstructured or semi-structured and are geographically distributed. Techniques have been developed to facilitate the cleaning, integration, aggregation and mining processes; also several peer to peer approaches have been developed to facilitate integration of health records geographical distribution, sequence labeling algorithms such as Conditional Random Fields (CRFs), Hidden Markov Models (HMMs), Artificial Neural Networks (ANN) are examples of techniques designed to handle the extraction of relevant information from unstructured health record [47]. For aggregation and data mining of both heterogeneous health records and handling of both structured and structured health records, several probabilistic and machine language have been developed; these include clustering algorithms such as SK-Means algorithm, Vector Space Model (VSM), Term Frequency–Inverse Document Frequency (TF-IDF) and natural language processing based techniques such as Doc2Vec and Word2Vec both used to create vectors from documents [20, 55]. An example of a developed solution is Decanter AI by Mobagel, capable of handling the above mentioned tasks using semi-supervised machine learning techniques [28].

4.3 Health Records Modelling

Health records model is a conceptual representation of the various elements in a health record, the relationship between them and the rules. Health record modelling is the process of creating a health model. Current health models are developed and evaluated using machine learning techniques, here the health record is typically split into a training set and test set. The training set is used to train a model optimizing its parameters while the test set is used after training has been completed to evaluate the model [48]. Challenges in health record modelling can be attributed to high dimensionality, heterogeneity, lack of fixed standards, sparsity and systematic bias of health records [22, 57] but advancement in deep learning and artificial intelligence help in overcoming these challenges [30].

4.4 Health Records Analysis

Health Records Analysis is the extraction of useful information from health records, health records analysis has multiple facets and approaches often using health record models. Deep learning techniques [30] could be used in health Records Analysis without human input resulting also in the automatic detection of hidden patterns and relationships between features while creating predictors. Deep learning techniques such as recurrent neural networks (RNNs), autoencoder, deep reinforcement learning, and convolutional neural networks (CNNs) [30, 43], are currently being used in health record analysis for example CNNsoftmax an end-to-end deep similarity learning approach for predictive analysis uses patient representation learning and patient similarity learning uses convolutional neural networks (CNNs).

Deep learning techniques [30] are popular because the can -

• Analyse multiple data types.
• Their performance improves as data-sets increases [2].
• detect hidden patterns and relationships in features.
• Discover learning representations without human input [58].

Popular Deep learning architectures currently used in health records analysis such as deep patient [41] an unsupervised learning architecture used to extract the hierarchical features and patterns in health records, deepr [36] a CNN based architecture used in predicting risk a patient faces by extracting features from health records, Med2Vec deepr [10] a Multi Layer Perceptron (MLP) based architecture allows for the interpretation of learned representation eliminating the need for expert knowledge in Medicine, and Doctor AI [9] a learning order dependence for diagnosis prediction.

4.4 Health Records Analysis

Health records contain valuable information that are also sensitive; for this reason, handlers of such records must ensure that they are kept secured and that privacy is maintained. Disputes may arise as to who owns health records, does the record belong to the
patient, the healthcare organization, the healthcare insurance, the government or is it jointly owned by some of the key players or all of them. Different countries have assigned ownership differently hence one needs to ensure that before attempting to perform any analysis on health records one should ensure that all the necessary approvals are gotten. [56]

5.2 Security

Securing health records is very important so also is ensuring that network, hardware, operating system, web, cloud and any vulnerable point of attack in healthcare facilities. Data security has evolved from using encryption alone to the adoption of several techniques [12]. Cryptography in used to secure health records against unauthorized access to private information, a popular technique cryptography technique is Hill cipher [38]. Two-factor authentication protocol is also an efficient and secure access protocol developed and adopted in the healthcare sector [49]. Also Blockchain is a decentralized technology that can be used as the foundation of a trustful and immutable system. The potential of blockchain has tested in the handling of health records [32]. Selecting the most effective blockchain model for the health system is however where the challenge currently lies as there are several blockchain-based applications, such as electronic medical record (EMR), MEDREC, MediLedger, Healthcoin [24].

Various techniques to ensure data security such as techniques that could support rural healthcare systems [18], detect abnormal behaviour [61] have been developed. Securing health records has received a lot of focus and techniques that suit almost all situations have been developed but the need for further research to mitigate against future challenges exist as new means to breach existing systems are constantly being explored.

5.3 Privacy

Maintaining the privacy of a patient’s record is important and often a significant concern to patients [25]. Health records contain sensitive information the patients often want kept private and if exposed could carry social humiliation, identity theft, economic impact, among others. Several techniques have therefore been developed to protect patients’ identities; some of the most commonly used techniques include -

- De-personalization - Removing as much personal identifiers from a patient medical record [42].
- Anonymization - The use of techniques such as suppression, generalization and slicing [31].
- Pseudonymization - Replacing patient identification data is replaced with an identifier.

5.4 Governance

Government and health boards including the World Health Organization all over the world have established various standards, guidelines or set of practices health workers are expected to follow especially when making critical decisions as these decisions could have significant negative effect. Below are some of the more popular health standards-

- General Data Protection Regulation (GDPR) in Europe [1, 47].
- Health Insurance Portability and Accountability Act (HIPAA) in the United States (US) [25].
- My Health Record (MHR) in Australia [23, 37].
- Electronic Health Record Architecture and Data Standard in China [60].
- eHealth-law, a "law on safe communication and applications in healthcare” In Germany [26].
- Health Information Technology for Economic and Clinical Health Act (HITECH) [29].
- Patient Protection and Affordable Care Act (PPACA) [29].

Governance also includes monitoring the health records and how they are processed to ensure that desired outcomes are meet; these are often done in form of an audit.

5.5 Sharing

The need to share health records across clinics, hospitals, pharmacies, laboratories, researches among others often arises [3], hence the need to ensure that those authorized to access and handle these records do so in a confidential manner following strict guidelines or standards. Without established standards or guidelines for sharing of health records between health facilities, these health facilities find it difficult to share data due to their need to ensure that the privacy of their patient is protected. A study estimated that 78 billion dollars could be saved annually if data exchange standards were utilized across board [29]. For example the European union (EU) understanding the diversities and peculiarities of member states, issued directives allowing member states via coordinated actions to expand the use of interoperable Electronic health records under its eHealth Action Plan 2012-2020 [26, 46].

5.6 Cost/Operational Expenditures

With health records constantly on the increase, storage and processing capacities need to be improved to handle the amount of data generated in an effective manner. Acquiring, setting up and managing storage and processing equipment that can handle the huge amount of data produced in health centers is expensive. Researchers and managers of health care facilities are constantly looking at ways to reduce the cost of managing such data centers or to seek for funding from other sources such as state or national government, non-governmental organizations and private establishments.

6 DEVELOPING NATIONS

There is no clear definition universally accepted for developing nations nor is there a clear agreement on which countries fit this category. In this study, we decided to classify countries with low Human Development Index (HDI) relative to other countries as developing nations. A questionnaire was developed carefully and given to health care professionals in developing countries to provide information on how generated data is stored, the quality of the data stored, and electronic sharing of patient’s data, among others.

Physicians and surgeons formed the major portion of those who responded to the questionnaire accounting for 79.6% of the responses, 11.8% was responses from nurses, nurse assistants and midwives while other categories such as pharmacists, records managers, dentists, optometrists, sonographers, healthcare administrator and general practitioner accounted for a little above a % each.
To understand the challenges hindering visualizing developing nations health records, the challenges involved in processing and managing health records were considered. Solutions to these challenges have been developed to some extent and researchers are actively seeking ways to further improve these solutions. The questionnaires that were issued to health care professionals in developing nations only focused on challenges that involve human input. Challenges hindering processing and managing health records -

6.1 Generation and Storage
To better understand how health records are captured and stored, questions posed to asked of health professionals investigated how they collected information they recorded in the health records, how these records are stored, and their ability to use electronic devices to capture information.

The results showed that manual storage of health records is still dominant, accounting for 45.2% of the total storage, electronic means account for just 10.8% while 44.1% use a combination of both means. An encouraging observation was that 76% of the respondents noted that their health facility is either planning to adopt electronic means to store records or had begun the process. Their responses also clearly showed that most health professionals understand how to use electronic devices to upload information effectively.

Dr Michael Kenechukwu Caled, a medical doctor in Nigeria, stated that “the cost and logistics involved in purchasing and setting up computers and other electronic devices in health facilities is the toughest step in transitioning from manual to electronic means. Should researchers develop tools that would enable health professionals use mobile phones and existing devices to electronically store health records, the adoption of electronic means would greatly improve”. Researchers also need to create tools that can capture non digital results such as Xray’s, manually documented report among others, so that even though records are generated manually, they can still be transferred to electronic forms.

6.2 Records cleaning, Integration and Aggregation
To understand the quality of the data and if they were integrated properly certain questions were asked.

It was observed that the quality of the health records is questionable as 88% of these workers have encountered situations where the records were in such poor states that they could not be used. It was also observed that less than 8% of the health professionals store information in the cloud, and less than 17% of the stored information are shared electronically. This implies that these records need to be cleaned to improve their quality. Integrating and aggregating health records currently in developing nations is therefore still very much a challenge.

Developing techniques to improve the accuracy of health records may need further research as syntax and semantics of health professionals in developing nations may differ from those used by those in developed nations. Natural language models like that of Bert and Elmo which were recently developed can be tested to see if they can be used to further improve the accuracy of health records.

6.3 Records Modelling
Health records modeling utilized techniques that could easily be transferred, and which have no unique characteristics for identification. Hence, it was not further investigated.

6.4 Records Analysis
Similar to health records modelling, health records analysis was not further explored.

6.5 Records Interpretation
Though there are several ways to present information obtained from analyzing health records for interpretation, using visualization techniques have proven to be the most effective. Developing nations may have difficulty affording powerful devices that could handle the processing of these visualization hence researchers need to proffer solutions that would allow developing nations use their existing devices and still visualize these health records.
6.6 Management Challenges

There are several management challenges hindering health records visualization in developing nations. Though these challenges are not unique to developing nations, developed economies have been taking steps to address these challenges. Developing nations need to follow suit. Challenges we suggest developing nations need to address for adequate visualization of health records are:

- **Ownership** - Health care organizations need to be permitted to analyze patients’ health records even without their consent in a manner that patients' privacy is protected. To achieve this existing standards need to be modified to encourage research while ensuring that patients privacy is protected.

- **Sharing** - With little or no sharing of health records via electronic means among developing nations, the need to encourage the use of already developed techniques that allow for easy sharing of information needs to be explored.

- **Cost/Operational Expenditure** - Just as in developed nations, health centers should be support financially to transition from manual storage to electronic methods.

- **Governance** - The number of health professionals that don’t follow strictly the guidelines as seen in figure 4 is too high. Authorities need to find ways of encouraging them to follow established standards and guidelines.

7 CONCLUSION

The process of extracting knowledge from data in the last decade has witnessed transformation as a result of successes in machine learning algorithms and artificial intelligence in almost all sectors including the health care sector. Visualization techniques have been used to present the information in a manner that allows for efficient extraction of knowledge or actionable information. In this survey paper, we gave a brief overview of health records, its characteristics and both processing and management challenges. We then showed that technological advancements have not improved the way knowledge is extracted in developing nations and suggestions were made to remedy the situation as the current global pandemic has shown that a health crisis in one nation could impact another or other nations.

REFERENCES

[1] Jan P. Albrecht. 2016. How the GDPR Will Change the World. *European Data Protection Law Review* 2, 3 (2016). https://doi.org/10.21552/EDPL/2016/3/4

[2] Esteve Andre, Robiquet Alexandre, Ramsundar Bharath, Kuleshov Volodymyr, A. DePristo Mark, Chou Katherine, Cui Claire, Corrado Greg, Thrun Sebastian, and Dean Jeff. 2019. A guide to deep learning in healthcare. *Nature Medicine* 25 (2019), 24–29.

[3] N Latha Anju, B. Murthy Rama, and U. Sunita. 2012. Electronic Health Record. *International journal of engineering research and technology* 1 (2012), 1–6.

[4] Aaron Bodd helicopters, William Harst, Michael Mackay, and Abdoumr El Rhahi. 2017. A Study into Data Analysis and Visualisation to Increase the Cyber-Resilience of Healthcare Infrastructures. In *Proceedings of the 1st International Conference on Internet of Things and Machine Learning* (LIV ’17). Association for Computing Machinery, New York, NY, USA, Article 32, 7 pages. https://doi.org/10.1145/3109761.3109793

[5] Jennifer Bresnick. 2017. Top 10 Challenges of Big Data Analytics in Healthcare. https://healthitanalytics.com/news/top-10-challenges-of-big-data-analytics-in-healthcare

[6] Luciano Caroprese, Pierangelo Veltri, Eugenio Vocaturo, and Ester Zumpano. 2018. Deep Learning Techniques for Electronic Health Record Analysis. In 2018 9th International Conference on Information, Intelligence, Systems and Applications (IISA). IEEE Computer Society, Zakynthos, Greece, 1–4. https://doi.org/10.1109/ISA.2018.8636467

[7] Valentina Casola, Aniello Castiglione, Kim-Kwang Raymond Choo, and Christian Esposito. 2016. Healthcare-Related Data in the Cloud: Challenges and Opportunities. *IEEE Cloud Computing* 3 (11 2016), 10–14. https://doi.org/10.1109/MC.2016.139

[8] Anitha Chinnaswamy, Armando Papa, Luca Dezi, and Alberto Mattiacci. 2018. Big data visualization, geographic information systems and decision making in healthcare management. *Management Decision* 57 (11 2018). https://doi.org/10.1108/MD-07-2018-0835

[9] Edward Choi, Mohammad T. Bahadori, Andy Schuetz, Walter F. Stewart, and Jimeng Sun. 2016. Doctor AI: Predicting Clinical Events via Recurrent Neural Networks. arXiv:1511.05942 [cs.LG]

[10] Edward Choi, Mohammad T. Bahadori, Elizabeth Searles, Catherine Coffey, Michael Thompson, James Bost, Javier Tejedor-Soto, and Jimeng Sun. 2016. Multi-Layer Representation Learning for Medical Concepts. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (San Francisco, California, USA) (KDD ’16) Association for Computing Machinery, New York, NY, USA, 1495–1504. https://doi.org/10.1145/2939672.2939823

[11] Alessandro Comai. 2014. Decision-Making Support: The Role of Data Visualization in Analyzing Complex Systems. *World Futures Review* 6, 4 (2014), 477–484. https://doi.org/10.1177/1946756715569233 arXiv:https://doi.org/10.1177/1946756715569233

[12] Kushal Dalal. 2019. A Novel Hybrid data security algorithm for Electronic Health Records security. (12 2019), 1–4. https://doi.org/10.1109/CSITSS47250.2019.9031012

[13] Declara H. David. 2017. What Does Big Data Look Like? Visualization Is Key for Humans. https://www.wired.com/insights/2014/01/big-data-look-like-visualization-key-humans/

[14] Reinse David, Gantz John, and Rydning John. 2017. The Evolution of Data to Visibility - Humans. https://www.wired.com/insights/2014/01/big-data-look-like-visualization-key-humans/

[15] Andrea De Mauro, Marco Greco, and Michele Grimaldi. 2016. A formal definition of Big Data based on its essential features. *Library Review* 65 (03 2016), 122–135. https://doi.org/10.1108/11108916500628614

[16] Dennis Egger, Edward Miguel, Shana S. Warren, Ashish Shenoy, Elliott Collins, Dean Karlan, Doug Parker, A. Mushifq Mobarak, Gunther Fink, Christopher Udry, Michael Walker, Johannes Hauskeller, Magdalena Larreboure, Susan Atthey, Paula Lopez-Pena, Salim Benachami, Macartan Humphreys, Layna Low, Niccolò F. Meriggi, Andrew Wahwire, C. Austin Davis, Ute Johanns Pape, Tilman Graff, Maarten Voors, Carolyn Nekesa, and Corey Vernot. 2021. Failing living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances* 7, 6 (2021). https://doi.org/10.1126/sciadv.abe0997 arXiv:https://advances.sciencemag.org/content/7/6/eabe0997.full.pdf

[17] RS Evans. 2016. Electronic Health Records: Then, Now, and in the Future. Yearbook of medical informatics. Yearbook of medical informatics

[18] Terzino. 2015. Big Data visualization, geographic information systems and decision making in healthcare management. *Management Decision* 57 (11 2018). https://doi.org/10.1108/MD-07-2018-0835

[19] RJ Evans. 2018. Deep Learning Techniques for Electronic Health Record Analysis. In *Proceedings of the 9th International Conference on Internet of Things and Machine Learning* (LIV ’17). Association for Computing Machinery, New York, NY, USA, Article 32, 7 pages. https://doi.org/10.1109/ISA.2018.8636467

[20] Agostino Forestiero and Giuseppe Papuzzo. 2020. Natural language processing approach for distributed health data management. 360–363. https://doi.org/10.1109/PDP50117.2020.00061
Electronic Health Records. BMC Medical Informatics and Decision Making 19 (12 2019), 259. https://doi.org/10.1186/s12911-019-0985-7

Paul Rubens. 2017. Big Data Visualization. https://www.datamation.com/big-data/big-data-visualization/

Abdul Samadzai, Vladimir Tomberg, and David Lamas. 2016. Current health records practices in Afghanistan and possible future development. 1–6. https://doi.org/10.1109/HealthCom.2016.7749431

Ezrim C. Schiza, Theodoros C Kyprianou, Nicolai Petkov, and Christos N Schizas. 2019. Proposal for an eHealth Based Ecosystem Serving National Healthcare. IEEE journal of biomedical and health informatics 23, 3 (May 2019), 1346–1357. https://doi.org/10.1109/jbhi.2018.2834230

Shahid Munir Shah and Burwan Ahmed Khan. 2020. Secondary Use of Electronic Health Record: Opportunities and Challenges. IEEE Access 8 (2020), 136947–136965. https://doi.org/10.1109/access.2020.3011099

F. Shamsot, T. Zhu, and D. A. Clifton. 2021. Machine Learning for Clinical Outcome Prediction. IEEE Reviews in Biomedical Engineering 14 (2021), 116–126. https://doi.org/10.1109/RBME.2020.3007816

M. I. Shahri, J. P. Li, S. Ullah, A. U. Haq, and G. Alam. 2019. An Efficient Access Privacy Protocol for Healthcare Patient Information System. (2019), 461–465. https://doi.org/10.1109/ICCWAMITP47768.2019.9067602

Uthayasankar Sivarajah, Muhammad Kamal, Zahir Irani, and Vinishth Weerakkody. 2016. Critical analysis of Big Data challenges and analytical methods. Journal of Business Research 70 (08 2016). https://doi.org/10.1016/j.jbusres.2016.08.001

Patil Sunil, Lu Hui, L. Saunders Catherine, Potoglou Dimitris, and Robinson Neil. 2015. Public preferences for electronic health data storage, access, and sharing - evidence from a pan-European survey. Journal of the American Medical Informatics Association 23 (06 2016). 1096–1098. https://doi.org/10.1109/jamiaoa.2016.9038969

Tableau. 2003. Data Visualization Tips For More Effective And Engaging Design. https://www.tableau.com/learn/articles/data-visualization-tips

Tableau. 2017. 10 Best Practices for Building Effective Dashboards. https://www.tableau.com/learn.whitpapers/10-best-practices-building-effective-dashboardsfor

Marko Teras and Shiraz Raghunathan. 2015. Big Data Visualisation in Immersive Virtual Reality Environments: Embodied Phenomenological Perspectives to Interaction. International Journal of Soft Computing 05 (08 2015), 1009–1015. https://doi.org/10.21917/ijsc.2015.0141

Shusaku Tsumoto and Shoji Hirano. 2014. Healthcare IT-Inetration of Consumer Healthcare Data and Electronic Medical Records for Chronic Disease Management. IEEE International Conference on Granular Computing (GrC) (2014).

Lidong Wang and Cheryl Alexander. 2019. Big Data Analytics in Healthcare Systems. International Journal of Mathematical, Engineering and Management Systems. 4 (02 2019), 17–26. https://doi.org/10.33889/IJMEMS.2019.4.1-002

Nicole Weiskopf, George Hirpsaca, Sushmita Swaminathan, and Chunhua Weng. 2013. Defining and measuring completeness of electronic health records for secondary use. Journal of biomedical informatics 46 (06 2013). https://doi.org/10.1016/j.jbi.2013.06.010

Cao Xiao, Edward Choi, and Jimeng Sun 2018. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. Journal of the American Medical Informatics Association 25, 10 (06 2018), 1419–1428. https://doi.org/10.1109/jamiaocp.2018.569498

Yu Ying, Li Min, Liu Liangliang, Li Yaohang, and Wang Jianxin. 2019. Clinical Big Data and Deep Learning Applications, Challenges, and Future Outlooks. BIG DATA MINING AND ANALYTICS 2, 12 (2018), 288–305. https://doi.org/10.26599/BDA/2019

C. Zhang, R. Ma, S. Sun, Y. Li, Y. Wang, and Z. Yan. 2019. Optimizing the electronic health records practices in Afghanistan. HealthCom 2016 Digital Media Industry Academic Forum (DMIAF) (2016). https://doi.org/10.1109/ACCESS.2016.2707584

Benjamin Goldstein, Ann Navar, Michael Pencina, and John Ioannidis. 2016. Opportunities of big data — A survey. (2017), 153–157. https://doi.org/10.1109/healthcom.2016.7749431

Eirini C Schiza, Theodoros C Kyprianou, Nicolai Petkov, and Christos N Schizas. 2019. Proposal for an eHealth Based Ecosystem Serving National Healthcare. IEEE journal of biomedical and health informatics 23, 3 (May 2019), 1346–1357. https://doi.org/10.1109/jbhi.2018.2834230

Shahid Munir Shah and Burwan Ahmed Khan. 2020. Secondary Use of Electronic Health Record: Opportunities and Challenges. IEEE Access 8 (2020), 136947–136965. https://doi.org/10.1109/access.2020.3011099

F. Shamsot, T. Zhu, and D. A. Clifton. 2021. Machine Learning for Clinical Outcome Prediction. IEEE Reviews in Biomedical Engineering 14 (2021), 116–126. https://doi.org/10.1109/RBME.2020.3007816

M. I. Shahri, J. P. Li, S. Ullah, A. U. Haq, and G. Alam. 2019. An Efficient Access Privacy Protocol for Healthcare Patient Information System. (2019), 461–465. https://doi.org/10.1109/ICCWAMITP47768.2019.9067602

Uthayasankar Sivarajah, Muhammad Kamal, Zahir Irani, and Vinishth Weerakkody. 2016. Critical analysis of Big Data challenges and analytical methods. Journal of Business Research 70 (08 2016). https://doi.org/10.1016/j.jbusres.2016.08.001

Patil Sunil, Lu Hui, L. Saunders Catherine, Potoglou Dimitris, and Robinson Neil. 2015. Public preferences for electronic health data storage, access, and sharing - evidence from a pan-European survey. Journal of the American Medical Informatics Association 23 (06 2016). 1096–1098. https://doi.org/10.1109/jamiaoa.2016.9038969

Tableau. 2003. Data Visualization Tips For More Effective And Engaging Design. https://www.tableau.com/learn/articles/data-visualization-tips

Tableau. 2017. 10 Best Practices for Building Effective Dashboards. https://www.tableau.com/learn.whitpapers/10-best-practices-building-effective-dashboardsfor

Marko Teras and Shiraz Raghunathan. 2015. Big Data Visualisation in Immersive Virtual Reality Environments: Embodied Phenomenological Perspectives to Interaction. International Journal of Soft Computing 05 (08 2015), 1009–1015. https://doi.org/10.21917/ijsc.2015.0141

Shusaku Tsumoto and Shoji Hirano. 2014. Healthcare IT-Inetration of Consumer Healthcare Data and Electronic Medical Records for Chronic Disease Management. IEEE International Conference on Granular Computing (GrC) (2014).

Lidong Wang and Cheryl Alexander. 2019. Big Data Analytics in Healthcare Systems. International Journal of Mathematical, Engineering and Management Systems. 4 (02 2019), 17–26. https://doi.org/10.33889/IJMEMS.2019.4.1-002

Nicole Weiskopf, George Hirpsaca, Sushmita Swaminathan, and Chunhua Weng. 2013. Defining and measuring completeness of electronic health records for secondary use. Journal of biomedical informatics 46 (06 2013). https://doi.org/10.1016/j.jbi.2013.06.010

Cao Xiao, Edward Choi, and Jimeng Sun 2018. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. Journal of the American Medical Informatics Association 25, 10 (06 2018), 1419–1428. https://doi.org/10.1109/jamiaocp.2018.569498

Yu Ying, Li Min, Liu Liangliang, Li Yaohang, and Wang Jianxin. 2019. Clinical Big Data and Deep Learning Applications, Challenges, and Future Outlooks. BIG DATA MINING AND ANALYTICS 2, 12 (2018), 288–305. https://doi.org/10.26599/BDA/2019

C. Zhang, R. Ma, S. Sun, Y. Li, Y. Wang, and Z. Yan. 2019. Optimizing the electronic health records through big data analytics: A knowledge-based view. IEEE Access 7 (September 2019), 136223–136231. http://eprints.whiterose.ac.uk/15695/1/ © 2019 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/.

Shengyao Zhou, Jie He, Hui Yang, Donghua Chen, and Runtong Zhang. 2020. Big-Data-Driven Abnormal Behavior Detection in Healthcare Based on Association Rules. IEEE Access 8 (07 2020), 1–1. https://doi.org/10.1109/access.2020.3099006

Ali Kalantari, Amirrudin Kamsin, Halim Kamaruddin, Nader Ale Ebrahim, Ab-