Laws and Regulations on Big Data Management: The Case of South Africa

Patrick Sello[1], Antoine Bagula[2(✉)], and Olasupo Ajayi[2]

[1] Department of Information Systems, University of the Western Cape, Cape Town 7535, South Africa
[2] Department of Computer Science, University of the Western Cape, Cape Town 7535, South Africa
abagula@uwc.ac.za

Abstract. A growing global trend has been witnessed in many developing countries where efforts and resources are been invested in advancement of electronic health information. The expectation is to improve the quality of health care, increase universal health coverage, and reduce both Legal Cases and healthcare costs in a changing world where data collected while providing healthcare produces big data sets which can provide useful insights for the advancement of healthcare services. The challenge is a greater risk for legal regulations to keep up with the accelerated global changes resulting from Big Data, and loss of information privacy created by digital transformation. In some countries, legal, privacy and ethical issues related to use and access to personal health data still causes foreseeable challenges. This article reviews the South African laws and regulations in handling, processing, storing, accessing and big data analytics on digital health data.

Keywords: Big data management · Laws · Regulations · Healthcare

1 Introduction

The digital revolution has changed how modern medicine is practiced as the use of information technology in healthcare delivery has grown rapidly in recent years. With it, volumes of digital health data are generated; which in turn improves the delivery of healthcare services, helps to address easy access to public healthcare, reduces information duplications and challenges faced by health professionals. A digital health record is the digital version of the patient's health record. Digital records are becoming common practice as more digital records are being created [2]. The creation of digital data records can assist in evidence based medical practice [4]. Privacy of information collected during healthcare processes is necessary because of the sensitivity content, stipulation of various legislations and protection of the patient's identity. With growing demand on the need for remote consolidation of digital health data, big data analytics, artificial intelligence and machine learning, the current legislative frameworks are no

R. Zitouni et al. (Eds.): AFRICOMM 2019, LNICST 311, pp. 169–179, 2020.
https://doi.org/10.1007/978-3-030-41593-8_12

longer able to cover and protect patient privacy. Most of the machine resources required for big data analytic, machine learning and artificial intelligence are only available in Cloud computing. However, this requirement for Cloud computing has encounter resistance within the public health sector due to legislative restriction on where and how digital health data can be stored and analysed. The management and access to digital health data requires [5]:

- Protection of information security, confidentiality and patient privacy always.
- Promoting information governance consensus among all stakeholders to use information better.
- Getting the basics right in terms of infrastructure, connectivity, basic ICT literacy, human resources and affordability planning.
- Taking an incremental approach.
- Adhering to the NHIS/SA principles for information management

Applicable security technologies exist and have proved effective in the banking and military sectors but experience is lacking to ascertain whether current technologies are satisfactory for health care. As yet, no model security implementations exist in any clinical computing environment, [6] although awareness of risks and of possible technical solutions is increasing [7]. This article will review the South African laws and regulations in handling, processing, storing, accessing and big data analytics on digital health data. The rest of this paper is arranged as follows, in Sect. 2 regulations for collection of digital health records are discussed, while Sect. 3 discusses regulations for processing of digital health records. In Sect. 4, focuses on storage of digital health records. Sections 5 and 6 focus on data management and access to digital healthcare data respectively. Section 7 discusses big data analytics in healthcare, while Sect. 8 focuses on integration and interoperability. Conclusion and potential future works are discussed in Sect. 9.

2 Digital Health Data Creation and Regulations

The benefits of digital health record include providing accurate, up-to-date, and complete information about patients at the point of care [2]. This:

- Enables quick access to patient records for more coordinated and efficient care
- Allows for sharing electronic information with patients and other clinicians in a secure manner
- Helps providers diagnose patients more effectively, thereby reducing medical errors.
- Improves patient-provider interaction and communication, as well as health care convenience
- Enables safer and possibly more reliable drug prescription.
- Helps promote legible, accurate and complete documentation as well as streamlined coding and billing

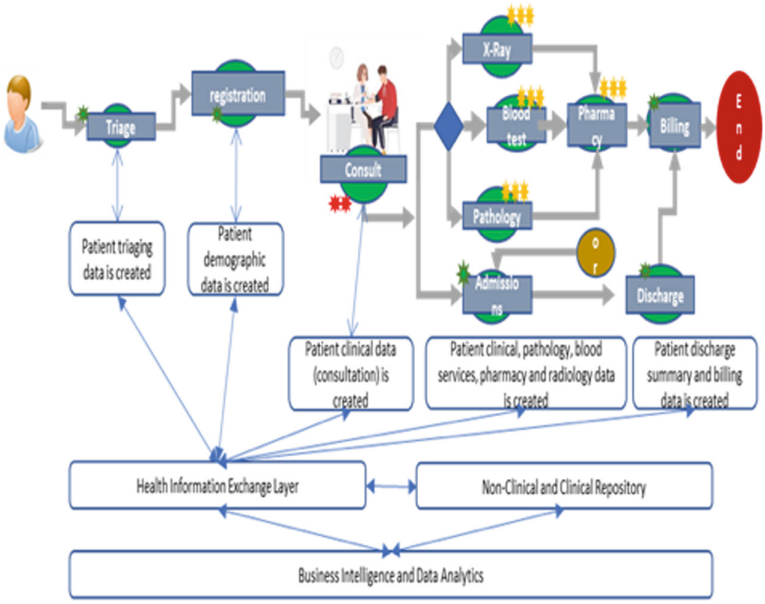

Fig. 1. Patient visit at academic hospital

- Enhances privacy and security of patient data
- Helps providers improve productivity and work-life balance
- Enables providers improve efficiency and meet their business goals
- Reduces costs by decreasing paperwork and duplicated tests.

Digital health data is created at every point of care within the patient journey in a hospital. Figure 1 depicts a high-level process flow of a patient's visit to an academic hospital.

The figure shows that data is created at triage, patient registration, consultation, radiology, pathology, admission and at the pharmacy. Data created in triage mostly comprises of patients' vitals which forms part of the clinical record. This data is collected by the healthcare professional. A Health Information System (HIS), is used by the healthcare professional to create, capture and collect digital health data during consultation with the patient. The system is used for patient administration, billing and collecting clinical data. The National Health Act 61 of 2003 Sect. 2, stipulate that, no health services can be provided to a patient without informed consent of the patient, unless the patient is unable to provide an informed consent. According to the Protection of Personal Information Act, 2013, a consent is defined as "any voluntary, specific and informed expression of will in terms of which permission is given for the processing of personal information" [10]. The Act is clear on the right of the patient being required for the data to be created, captured and collected. Patient demographics data is collected during registration, this data is used to identify the patient, provide next of kin, verify the economic status of the patient or the person responsible for

the payment of the services, to be rendered. The registration process to capture, create and collect digital health data is carried-out by hospital clerks or administrators. Section 14 of the National Health Act stipulates that "all information concerning a patient, including information relating to his or her health status, treatment or stay in a health establishment, is confidential" [7]. However, the Act allows for the following exceptions to this general rule: (a) when the user consents to that disclosure in writing; (b) when a court order or any law requires that disclosure; or (c) when the non-disclosure of the information represents a serious threat to public health. The Act is limited on processes for vetting people who are capturing, creating and collecting digital health data. This has led to a wide range of abuse and misuse of digital health data by academic institutions, NGO's, NPO's, government departments and other independent agents. As an example, an international NGO which had a partnership with Gauteng Department of Health (GDoH), was allowed to use its own employees to assist the department in capturing, creating and collecting digital health data of patients with chronic diseases. It was later discovered that, the NGO used most the data for research on behalf of international pharmaceutical companies. The NGO was rendering a legitimate service as agreed with GDoH, however, because of shortfalls within the legislative framework it is a difficult process to provide vetting for people responsible for collecting digital health data. The confusion and vulnerability is further exacerbated by the same National Health Act, which also states that, "a healthcare professional or healthcare worker that has access to the health records of a patient can disclose such personal information to any other healthcare professional as is necessary for any legitimate purpose within the ordinary course and scope of his or her duties where such access or disclosure is in the interests of the patient." This make digital health data privacy difficult to police and monitor adherence regulations [8].

3 Digital Health Data Ingestion (Processing) and Regulations

The European Union's General Data Protection Regulation (GDPR) states that, "higher protection standards for health data and delineates a variety of definitions and conditions that apply to such data," it highlights the required conditions that must be meet for health data to be processed [11]. The conditions include explicit consent of the patient, clear purpose of why the data must be processed and public interest. The ingestion of digital health data starts as soon as the data is collected using the Health Information System (HIS). Health information technology apply information processing tools both hardware and software to the storage, retrieval, sharing, and use of health care information, data, and knowledge for communication and decision making [13]. Health Information Systems thus facilitate the process of storing and retrieval of individual records with the aid of computers interconnected through a computer network. Once the digital health data is captured (as described in the previous section), a patient unique identifier is generated and allocated to the patient. The patient

unique identifier becomes the Master Patient Index (MPI) which forms part of Enterprise Master Patient Index (EMPI). An EMPI is a patient database used by Healthcare facilities to maintain current and accurate digital health data across multiple healthcare systems. The allocated patient unique identifier is presented only once across all healthcare systems. South African National Department of Health has recently launched a Health Patient Register System (HPRS), the system is aimed to process digital health data and create single patient identifiers across the country which will allow patient to receive healthcare services across provinces. Once the data has been processed, there are multiple storage platforms which are in use and available for the data to be stored. With respect to storage, the Protection of Personal Information Act 2013, stipulate that, the responsible party must ensure that the conditions set out in Chap. 3 condition 1 of the Act is met "and all the measures that give effect to such conditions, are complied with at the time of the determination of the purpose and means of the processing and during the processing itself." The Act goes further to stipulate that, "the data must be processed lawfully, in a reasonable manner that does not infringe the privacy of the data subject and only be processed if, given the purpose for which it is processed, it is adequate, relevant and not excessive". Section 2 of the National Health Act covers only research related data, there is no clear guidelines or regulated related to how general digital health data must be processed. At the international level, systems for surveillance and monitoring of diseases and epidemics, and initiatives to share knowledge and data for health research and health development are progressing. WHO continues to track the adoption of eHealth and Universal Health Coverage (UHC) goals and measuring the achievement of Sustainable Development Goals (SDGs). Two particular work areas continue to advance eHealth agenda in this respect: the Health Data Collaborative and the implementation of the WHO Framework for Integrated People-Centred Services [15]. In some countries there is still a need to build a strong eHealth foundation including necessary infrastructure, standards, legislation and workforce. Rules and regulation are very limited on how digital health data can be processed and the legality of how digital health data must be processed. Legal, privacy and ethical issues related to use and access to personal health data still causes foreseeable challenges in many countries. However,"WHO is expanding its focus on digital health, the Organisation has been working in this area for years, for example, through the development of the eHealth Strategy Toolkit in 2012, published in collaboration with International Telecommunications Union (ITU)" [3]. Because of these gaps within the legislative framework, academic institutions across the world have found ways of processing digital health data for research purpose which is collected from public healthcare facilities, while absolving themselves of any form of responsibility as stipulated in the Protection of Personal Information Act.

4 Storage of Digital Data

The National Health Act, stipulate that, "Subject to National Archives of South Africa Act, 1996 (Act No. 43 of 1996), and the Promotion of Access to Informa-

tion Act, 2000 (Act No. 2 of 2000), the person in charge of a health establishment must ensure that a health record containing such information as may be prescribed is created and maintained at that health establishment for every user of health services."

Figure 2 depicts multiple stage where digital health data is stored. The figure provides a high-level view of were the data is generated versus stored. All the data generated during consultation with the patient is a relational database linked to a module within the HIS. Different types of digital health data are created at each stage, the data is later stored in a clinical repository which provides access to data analytic, big data and data visualization. Condition 3 Section 14 of the Protection of Personal Information Act stipulate that, "Subject to subsections (2) and (3), records of personal information must not be retained any longer than is necessary for achieving the purpose for which the information was collected or subsequently processed, unless;

(a) retention of the record is required or authorised by law;
(b) the responsible party reasonably requires the record for lawful purposes related to its functions or activities;
(c) retention of the record is required by a contract between the parties thereto; or
(d) the data subject or a competent person where the data subject is a child has consented to the retention of the record".

This is in support of the National Health Act, which provides guidelines on how long health records can be kept in a healthcare facility. The Health Professional Council of South Africa (PCSA) offers the following guidance on the retention of medical records: (a) Records must be stored for 6 year after becoming dormant. (b) Records of people under 21 must be kept until they reach 21 years. (c) Mental health records must be kept until the death of the patient receiving treatment [17]. These regulations and guidelines were formulated on a paper-based records approach. These records had a limited time span, were prone to damage and required huge warehouse storage. Unlike paper-based records, digital health data can be stored and archived for many years in digital warehouses, data centres, Cloud storage and on-premise storage devices. Digital data can also be stored anywhere in the world if there is adequate connectivity, however, this creates a gap on legislative requirements on how and where the data is stored, and the duration for which such data can be stored. The legislative framework requires digital health data to be stored where the patient is receiving treatment. The Protection of Personal Information Act stipulate that, "no personal information which includes digital health data can be transferred outside the Republic unless, the third party who is the recipient of the information is subject to a law, binding corporate rules or binding agreement which provide an adequate level of protection". This means that, data cannot be shared outside the healthcare facility where the patient is receiving treatment. It also means that any form of international collaboration will not be possible. The major limitation to the framework is that, with the growing demand for more computer resources for big data analytics, machine learning and artificial intelligence, which Cloud

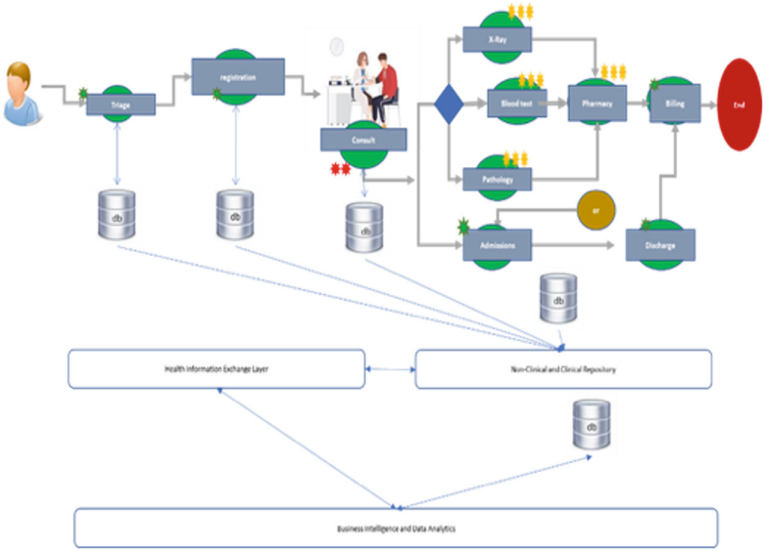

Fig. 2. Digital health data

computer can provide. Unfortunately, these computing power are hosted in first world developed countries, and because of legislative limitations digital health data cannot cross country boarders. This limitation is due to the failure of policy makers to accelerate digital transformation in public healthcare facilities. A gap has been however been created for academic intuitions to extract and collect digital health data. The approach is to provide an electronic health record system which is used by intern doctors to capture data related to patients, the intern doctors are incentivized by being able to use the collected data as portfolio of evidence(s) for their community service.

5 Data Management

Data integrity and quality of the data collected must be of high standard at all time. The data contains information about the medical history of the patient which must be accurate and updated whenever there is an encounter with the patient. The Protection of Personal Information Act provides legal guidelines on steps to ensure that the personal information is complete, accurate, not misleading and updated where necessary. In taking these steps, the responsible party must have regard to the purpose for which personal information is collected or further processed. Data Ownership is a legal and regulatory complex discussion. The health information is owned by the patient as directly prescribed in the Protection of Personal Information Act and Protection of Access to Information Act. The discussion has always been on the person responsible for creating the record, or where the record was created as been the custodians of the record.

"In the case of public health institutions, where records e.g. radiographs are the property of the institution, original records and images should be retained by the institution. Copies must however, be made available to the patient (or referring practitioner) on request for which a reasonable fee may be charged in terms of the Promotion of Access to Information Act (Act No. 2 of 2000)" Data Security is one of the difficult aspects associated with the creation of digital health data. The Nation Health Act stipulate that, "the person in charge of a health establishment in possession of a user's health records must set up control measures to prevent unauthorised access to those records and to the storage facility in which, or system by which, records are kept". The Protection of Person Information Act also puts the responsibility of securing personal records on the person responsible for the collecting the information. It further states that, reasonable technical and organisational measures must be put in place to prevent: (a) loss of, damage to or unauthorised destruction of personal information; and (b) unlawful access to or processing of personal information."

This can be done by identifying all reasonably foreseeable internal and external risks to personal information stored within the facility or organisation. The responsible party must have due regard to generally accepted information security practices and procedures which may apply to it generally or be required in terms of specific industry or professional rules and regulations.

6 Access to Digital Health Data

The Health Act states that, "no health care practitioner shall make information available to any third party without the written authorisation of the patient or a court order or where non-disclosure of the information would represent a serious threat to public health". This Act has an inherent limitation as it allows for health data to be disclose during legal matters, where non-disclosure of the medical information has a threat to public health and/or where it is in the interest of the wellbeing of the patient and there is a health risk to the patient if the information is not disclosed. Measures must therefore be put in place that will provide audit trail of any other person(s) who has access to the information and the purpose linked to that access.

7 Access for Big Data and Analytics

The introduction of Big Data has revolutionized how we manage, analyse and take advantage of data to provide healthcare and management decision making. Healthcare is one of the promising early gains in the use of big data for management decision making and planning. As data is collected within multiple healthcare systems in high volume, high velocity and high variety, it becomes para-important for deploying mechanisms for mining and analysing data. As an example, Discovery Holdings, a company which specialises in health and life insurance, has been using big data technology to promote healthier behaviour. The approach has been studied to produce the required outcomes base on a study

conducted by RAND Europe [18]. There are couple of technological platforms such as Hadoop and Spark which can be can be used in the public healthcare for predictive analysis, real-time patient monitoring (with IoT), medical equipment monitoring (with IoT), Electronic Health Records (EHRs), genomics, Tracking of communicable disease in boarder crossing, reduction of re-admissions and building artificial intelligence models for early detection of cancers and other disease, etc. Most governments have not progressed in reshaping national policies to improve the use of big data while adhering to internal regulations on data privacy, confidentiality and security. According to the WHO's Global Observatory for eHealth, only 21 (17%) of the 125 Member States reported having a policy or strategy regulating the use of big data in their health sectors [19]. The responsibility of creating progressive legislation on big data has been left to academic institutions whose interest does not align with government policies on universal health coverage and digital health. However, even basic health data can be misused, potentially leading to discrimination, especially of the vulnerable populace [20].

8 Integration and Interoperability

There can be over 200 parties in a standard public healthcare environment, this creates standardisation challenges in the system and results in most systems and data being fragmented. As depicted in Fig. 1, data is created in multiple points and collected across systems which are incompatible to each other. The business process of generating such data are also not aligned and imputable. The Health Normative Standards provides a set of standards-based profiles which must guide any interoperability function within healthcare. It has been shown in multiple studies that, poor coordination of technology and lack of standards are limiting factors for facilitate collaboration. Standardisation of business processes within healthcare can assist in addressing such challenges. The traditional based regulations are limited when covering integration regulations.

9 Conclusion and Future Work

The continues increase of technology within the healthcare will require coordination across all stakeholders to ensure that benefits to healthcare are truly achieved [20]. Policy makers will have to review and make necessary amendments to realise full benefit of big data and data collaborations in healthcare. The answer lays in the interaction between institutions of higher learning and policy makers within government. There must be a strategy to coheres exiting academia knowledge that can be used to accelerate amendments of policies related to digital data privacy and collaboration analytics. The focus must be on getting the basics right, taking incremental approach, looking for early wins and advocating the benefits to healthcare. As an avenue for future research, the work presented in this paper in the South African context can be adapted to guiding policy in the implementation of cyberhealthcare systems [21–26] in both rural and urban areas of the world and especially in developing countries.

References

1. Barrows, R.C., Clayton, P.D.: Privacy, confidentiality, and electronic medical records. J. Am. Med. Inf. Assoc. **3**(2), 139–148 (1996). https://doi.org/10.1136/jamia.1996.96236282
2. Rouse, M.: "What is electronic health record (EHR)? - Definition from WhatIs.com", SearchHealthIT. https://searchhealthit.techtarget.com/definition/electronic-health-record-EHR. Accessed 30 May 2019
3. "What are the advantages of electronic health records? — HealthIT.gov", Healthit.gov, 2019. https://www.healthit.gov/faq/what-are-advantages-electronic-health-records. Accessed 30 May 2019
4. South African Minister of Health, eHealth Strategy, p. 8 (2012)
5. Shea, S.: Security versus access: trade-offs are only part of the story. JAMIA **1**, 314–5 (1994)
6. Barrows, R., Clayton, P.: Privacy, confidentiality, and electronic medical records. J. Am. Med. Inf. Assoc. **3**(2), 139–148 (1996). https://doi.org/10.1136/jamia.1996.96236282
7. National Health Act 61 of 2003
8. Promotion of Information Act 2 of 2000
9. Agarwal, T.K.: Vendor neutral archive in PACS. Indian J. Radiol. Imag. **22**(4), 242–245 (2012). https://doi.org/10.4103/0971-3026.111468. Accessed 1 June 2019
10. Cognizant 20–20 Insights, "The U.S. Healthcare Implications of Europe's Stricter Data Privacy Regulations (2018)
11. Almunawar, M., Anshari, M.: Health Information Systems (HIS): concept and Technology (2012). Accessed 2 June 2019
12. Protect of Personal Information Act 2013
13. Geneva, 14 to 18 May 2018, Commission on science and technology for development (CSTD) (2018)
14. Purtova, N., Kosta, E., Koops, B.-J.: Laws and regulations for digital health. In: Fricker, S.A., Thümmler, C., Gavras, A. (eds.) Requirements Engineering for Digital Health, pp. 47–74. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-09798-5_3
15. Chaib, F., Garwood, P.: WHO releases first guideline on digital health interventions (2017). https://www.who.int/news-room/detail/17-04-2019-who-releases-first-guideline-on-digital-health-interventions. Accessed 27 May 2019
16. HPCSA, Guidelines on the Keeping of Patient Records, paragraph 9 (2008)
17. Marr, B.: This Health Insurance Company Tracks Customers' Exercise And Eating Habits Using Big Data And IoT(2019). https://www.bernardmarr.com/default.asp?contentID=1884. Accessed 03 Jun 2019
18. World health organization., global diffusion of ehealth. World health organization, geneva (2017)
19. Vayena, E., Dzenowagis, J., Brownstein, J., Sheikh, A.: Policy implications of big data in the health sector. Bull. World Heal. Organ. **96**(1), 66–68 (2017). https://doi.org/10.2471/blt.17.197426
20. Taylor, K.: Digital health the future of healthcare — life sciences and healthcare — Deloitte Southern Africa. https://www2.deloitte.com/za/en/pages/life-sciences-and-healthcare/events/digital-health-the-future-of-healthcare.html. Accessed 26 May 2019

21. Mandava, M., Lubamba, C., Ismail, A., Bagula H., Bagula, A.: Cyber- healthcare for public healthcare in the developing world. In: proceedings of the 2016 IEEE Symposium on Computers and Communication (ISCC), Messina-Italy, 27–30 June 2016, pp. 14–19 (2016)
22. Bagula, M.F., Bagula, H., Mandava, M., Kakoko Lubamba, C., Bagula, A.: Cyber-healthcare kiosks for healthcare support in developing countries. In: Mendy, G., Ouya, S., Dioum, I., Thiaré, O. (eds.) AFRICOMM 2018. LNICST, vol. 275, pp. 185–198. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-16042-5_18
23. Celesti, A., et al.: How to develop IoT cloud e-health systems based on FIWARE: a lesson learnt. J. Sens. Actuator Netw. **8**(1), 7 (2019)
24. Bagula, A., Mandava, M., Bagula, H.: A framework for healthcare support in the rural and low income areas of the developing world. J. Netw. Comput. Appl. **120**, 17–29 (2018). https://doi.org/10.1016/j.jnca.2018.06.010
25. Bagula, A., Lubamba, C., Mandava, M., Bagula, H., Zennaro, M., Pietrosemoli, E.: Cloud based patient prioritization as service in public health care. In: 2016 ITU Kaleidoscope: ICTs for a Sustainable World (ITU WT), pp. 1–8. IEEE (2016)
26. Lubamba, C., Bagula, A.: Cyber-healthcare cloud computing interoperability using the HL7-CDA standard. In: 2017 IEEE Symposium on Computers and Communications (ISCC), pp. 105–110. IEEE (2017)