An Ethical Framework for Big Data and Smart Healthcare

Victor Chang[1][a], Rahman Olamide Eniola[2][b], Ben Shaw-Ching Liu[3][c] and Mitra Arami[4][a]

[1]*Department of Operations and Information Management, Aston Business School, Aston University, Birmingham, U.K.*

[2]*Cybersecurity, Information Systems and AI Research Group, School of Computing, Engineering and Digital Technologies, Teesside University, U.K.*

[3]*Department of Marketing, Lender School of Business Center, Quinnipiac University Hamden, CT 06518, U.S.A.*

[4]*Pardis Limited, London and EM Normandie Business School, France*

Keywords: Ethics for AI and Data Science, Ethical Framework, Ethics for Smart Healthcare.

Abstract: There has been significant growth in big data technology in healthcare in recent years. However, the potential of big data analytics is affected by various ethical and security concerns, which have hampered the application of big data analytics in healthcare. Recently, numerous studies have been conducted on the emerging big data ethical issues in healthcare. While most of the journal reflects on privacy and security questions, it did not examine; objectively the possible discriminatory impact of big data analytics has no. This mixed-method project aims to highlight various ethical problems in big data analytics while also providing an in-depth insight into the biased results derivable from big data analytics and the effects of such outcomes.

1 INTRODUCTION

Higher healthcare investment in a nation can provide better health prospects that can enhance human capital and increase productivity, thus contributing to economic performance (Raghupathi and Raghupathi, 2020; Cutillo et al., 2020). However, the exponential growth in the world's population presents a critical threat to current medical and healthcare systems (Zhu et al., 2019). The change in population demographics, the increase in the number of aged people, and the drastic increase in the cost of in-hospital services all lead to realizing the value of effective healthcare systems (Demirkan, 2013). The professional-to-patient ratio is another factor that led to the rise in demand for an efficient healthcare system (Borodin et al., 2016).

With the explosive growth of disruptive technologies in recent years, the speed and quantity of digital data collected have expanded steadily and rapidly (Chang, Shi and Zhang, 2019). Correspondingly, the evolution of information technology and the introduction of digitized computer systems has resulted in the transition of conventional hard copy medical data to Electronic Health Records (EHR) and Electronic Medical Records (EMR) systems (Rehman, Naz and Razzak, 2021). These systems resulted in exponential data expansion (Razzak, Imran and Xu, 2020), which has contributed to the growth of big data analytics, especially in healthcare.

According to a 2021 Grand View Research, Inc. study, the worldwide healthcare analytics market was valued at USD 23.6 billion in 2020, projected to rise at a Compound Annual Growth Rate (CAGR) of 23.8 percent from 2021 to 2028 ('Healthcare Analytics Market Size Industry Report: 2021-2028', 2021). See Figure1.

This massive increase fulfills the growing need for improved healthcare, aided by innovative technology.

Administrative claim reports, hospital registries, electronic records of health, biometric data, patient data, the internet, medical imaging, biomarkers, prospective cohort studies, and clinical trials are possible medical big data sources in healthcare (Hermon and Williams, 2014; Luo et al., 2016).

[a] https://orcid.org/0000-0002-8012-5852
[b] https://orcid.org/0000-0001-9799-861X
[c] https://orcid.org/0000-0002-2950-9607
[c] https://orcid.org/0000-0001-6855-9888

Chang, V., Eniola, R., Liu, B. and Arami, M.
An Ethical Framework for Big Data and Smart Healthcare.
DOI: 10.5220/0011030900003206
In *Proceedings of the 4th International Conference on Finance, Economics, Management and IT Business (FEMIB 2022)*, pages 65-74
ISBN: 978-989-758-567-8; ISSN: 2184-5891

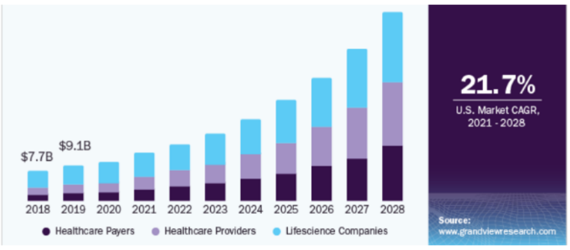

Figure 1: Healthcare Analytics market in the USA, by end user, 2018 – 2028 (USD Million).

These sources; are aggregated to produce fast and cost-effective prescriptive, descriptive, and diagnostic insights for the healthcare stakeholders. While strategically analyzing data for insightful analysis is crucial, the existence of different data types accessible from numerous sources makes big data management extremely difficult (Nair, 2020).

Despite the aforementioned benefits of big data technology, it is worth remembering that big data analytics has its drawbacks due to its intertwinement with people's sensitive personal information, daily behavioral patterns, and potential prospects. The most pressing concerns of big data analytics are privacy (Francis, 2014), confidentiality and informed consent (Ioannidis, 2013), epistemic hurdles (Floridi, 2012), and the analysis of monitoring in a growing datafication of the society (Ball et al., 2016). Indeed, the assurance of privacy and safety of subjects through the application of big-data analytics are of significant importance and high priority. The study by IBM (IBM, 2019) in Figure 2 shows that the health sector has suffered an average overall cost of data breaches considerably higher than other sectors such as hospitality, media, and research. Healthcare data should be securely kept, and big data analytics performed ethically (Mittelstadt, 2019).

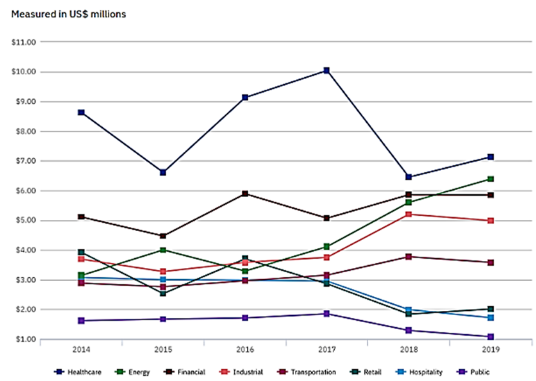

Figure 2: Trend of Industry Average Data Cost (IBM, 2019).

The possibility of potential discrimination is among the most alarming yet understudied issues of big data technology. There is no universally accepted definition of discrimination. The term generally refers to acts, practices, or policies that impose a relative disadvantage or treat a person or specific group of people differently, especially in a worse way than treating other people because of their skin color, gender, sexuality, language, or other factors (Reinsch and Goltz, 2016).

The research (Obermeyer et al., 2019) that revealed pervasive racism in decision-making systems utilized by US clinics is an excellent demonstration of discrimination in healthcare analytics. Participants who self-identified as black were rated lower risk scores than equally ill white people in the study. Consequently, black individuals were less likely to be referred for more personalized medical care (Obermeyer et al., 2019).

The emergence of these instances describes why discrimination in big data analytics has become an emerging topic in a variety of fields, from data science and artificial intelligence to psychology, culminating in a dispersed and fractured interdisciplinary corpus that tends to make thoroughly accessing the foundation of the problem difficult (Favaretto et al., 2019).

This study summarized big data and its use in healthcare, addressing current ethical and security issues relevant to big data application in healthcare. Moreover, we suggest several alternative solutions to compromise between the application and the ethical obligation.

2 LITERATURE REVIEW

Big data and big data analytics are arguably the pillars of other disruptive technologies, providing the necessary business insights for patients, experts, and government (Wong, Zhou, and Zhang, 2019). Big data analytics is the method of storing, processing, and analyzing vast collections of data to find trends and other valuable knowledge (Heyman et al., 2004). These massive and complex big data collections are manipulated and managed using various computational methods such as machine learning and artificial intelligence (Ward and Barker, 2013). The advent of advanced technology has provided conditions and procedures for voluminous databases to be compiled and processed, resulting in informed decision-making in addressing health problems (Raja et al., 2020).

Big data has emerged as a promising option with the potential to revolutionize the healthcare system by lowering costs and optimizing treatment process, delivery, and management (Patil and Seshadri, 2014). The application of big data comes with some ethical issues that demand careful consideration (Camilleri, 2020). Suresh and Guttag (2019) explain how bias problems occur, how they apply to specific applications, and how they inspire various solutions. They also present a framework for understanding analytical bias at a higher level of abstraction to facilitate constructive dialogue and solution development.

Notwithstanding the amount of data generated in healthcare, the underlying challenge remains in the integration, of structured and unstructured health data. According to Dridi et al. (2020), approximately 80% of clinical data is unstructured: and widely underutilized, once generated. Different clinical data formats, such as scanned canned medical documents, prescriptions, patient registries, and clinician notes, result in poor standardization of healthcare data, making it more difficult to handle by EHR systems and more prone to bias from data preprocessing (Cave et al., 2019; Dridi et al., 2020).

Patient privacy invasion is an emerging problem in big data analytics. Patients' behavior and sentiment data can be obtained from various online sources. For example, an online drug retailer may have recorded the purchase of a particular medication, a ride-hailing app may have recorded a visit to a clinic or lab, or a social media app may have recorded patients' interactions with a medical web page. Furthermore, patients' data can also be extracted unethically via health-care-specific applications and wearable devices.

Also, we studied several publications to grasp better the potential discriminatory effects and popular drivers of discrimination or inequality in big data analytics on subjects. Different writers arrived at different conclusions. Big data analytics may result in unintentional discrimination (Žliobaitė, 2017; Sonawane and Irabashetti, 2015). Žliobaitė (2017) established that discrimination is indirect, not by the analyst's intention but because of the structure and noise of experimental data. Such algorithms may systematically disfavor persons belonging to particular groups or categories, rather than depending purely on individual merits.

Conversely, other academic studies emphasized intentional discrimination (e.g., Kuempel, 2016; Sonawane and Irabashetti, 2015). According to Kuempel (2016), data brokers frequently combine raw components of personal data in a discriminatory way, leaving customers exposed to exploitative and distasteful marketing techniques. The effect of utilizing such a biased dataset with sensitive information is that such individuals or groups of people would lead to direct discrimination.

Suresh and Guttag (2019) explain how bias problems occur, how they apply to specific applications, and how they inspire various solutions. They also present a framework for understanding analytical bias at a higher level of abstraction to facilitate constructive dialogue and solution development.

3 RESEARCH QUESTIONS AND BIG DATA ANALYTICS ARCHITECTURE

The first step of the research was to identify relevant research questions. The main research question is, "given the many applications and benefits of big data and big data analytics in healthcare, do the ethical risks overshadow the benefits?"

To answer this main question, we need to find answers to the following sub-questions:

1. What are the applications of big data in healthcare?
2. What are the current ethical issues of healthcare big data analytics?
3. What is the cause of discrimination in big data analytics?

The big data analytics framework utilized in this project is a blend of many steps that explains the big data Analytics procedure (shown in Figure 3 above). The first phase in the framework is data preparation, which involves the ETL, i.e., Extraction, Transformation, and Loading of the data. Extraction is the process of determining the data type to be utilized and collecting it from different data sources, such as existing databases and repositories, APIs, and the cloud. Data transformation is the next step in which data is transformed, aggregated, and loaded into the Power Business Intelligence (BI) dashboard.

Figure 3: Power BI architecture.

The transformation step is to ensure the: (1) handling of inconsistencies and missing values in the data; (2) elimination of duplicate data; (3) removal of useless data; and (4) sorting of data into the appropriate type. Figure 3 below illustrates the overview of the Power BI analytics procedure.

The visualization step involves taking the processed outputs and transforming them into meaningful insights by viewing the results in diagrams, KPIs, or other easy-to-understand formats. It is crucial to ensure that results can be interpreted by those with no previous experience or expertise. Unlike other tools, Power BI allows the integration of different programming languages. Applying Python and R functionalities while using the DAX and M-language formulas is the advantage of Power BI. It gives a better result due to the combined strengths of different programming languages.

4 APPLICATION AND BENEFITS OF BIG DATA ANALYTICS

4.1 Preventive Medicine

Preventive medicine is arguably the most innovative application of big data analysis which employs cutting-edge data analytics methods: for disease detection and classification, association analytics, and clustering, with the promise of efficiently discovering valuable patterns by analyzing large amounts of unstructured, heterogeneous, non-standard data (Razzak et al., 2020). Appropriate disease prevention involves identifying and treating at-risk patients. To increase therapeutic adherence, several preventative strategies are employed. Pertinent data, such as body temperature, pulse, and blood pressure, are electronically collected, enabling automated risk prediction. Consonantly, the increased usage has contributed significantly to the appropriateness of big data analytics in healthcare (Rehman et al., 2021). The aggregate of these data is analyzed to assist patients with diets, reminders of preventative care, personalized medical care, follow-up on prior consultations and medicines, and counseling (Razzak et al., 2020).

Due to the considerably broad customer base, relatively few regulatory obligations and ease of access to wearable devices and medical apps, personalized medical care has significantly increased its market size, as shown in Figure 4.

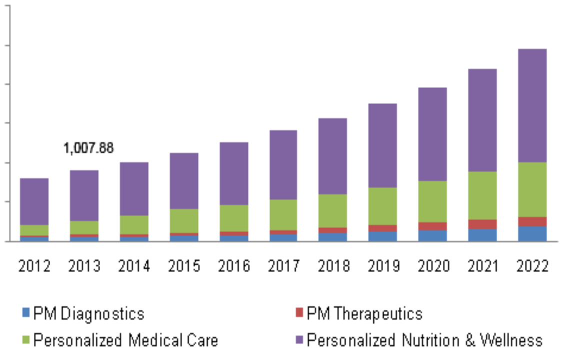

Figure 4: Trend of Personalized Medicine (2012 -2022).

4.2 Evidence-based Healthcare

Traditional healthcare is changing from expedient and discretionary decision-making to evidence-based medical practices (Piai and Claps, 2013).

Evidence care is a healthcare practice where we base the patients' conditions on scientific proof. Through consolidating data from various outlets, big data offers evidence-based treatment. The data trends and patterns would provide sufficient support for diagnosis and treatment (Piai and Claps, 2013).

4.3 Enhancement of Public Health Monitoring

The analysis of healthcare data with ground-breaking methods aids in the epidemic trends analysis, disease outbreaks monitoring, and the spread of disease. This approach improves public health monitoring, education, and reaction time. An excellent example is the Covid 19 pandemic surveillance system in the United Kingdom which offers a daily update of a postcode district-based location with infection rates in that district, generates a risk score, and communicates it to the user. Furthermore, the app allows users to check into a specific place, recording their presence at that particular time and date. The app also stores an individual's check-ins with the name and IDs of such locations, which work with the test and trace teams to inform users on association with a particular area at a given time. For example, suppose someone visits a local bar and is tested positive with Coronavirus. In that case, the app alerts everyone who has also checked in the same place to self-isolate or quarantine.

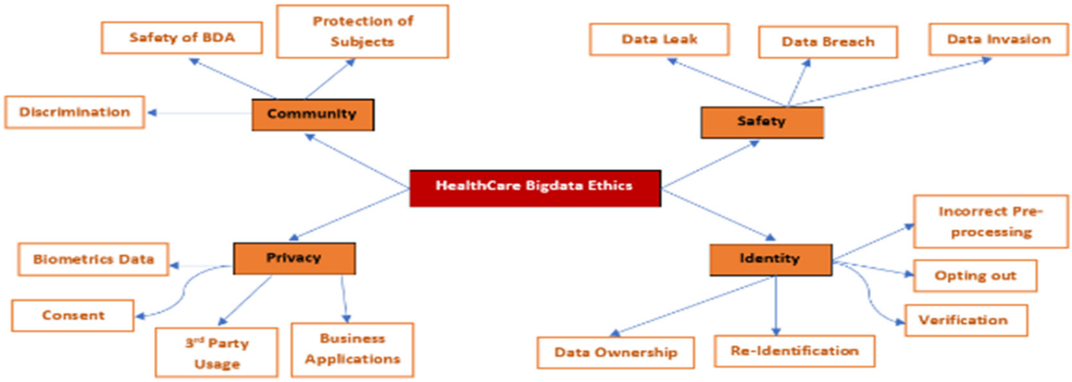

Figure 5: Emerging Ethical Issues in Healthcare Big data.

4.4 Improves Interaction between Healthcare Providers and Patients

Big data technology also improves collaboration between healthcare providers and patients. For example, on social media, people with common health conditions and healthcare professionals with similar specialties across the world can share information on the treatment and cure of some illnesses, thereby promoting interaction within health systems.

5 METHODOLOGY

5.1 Systematic Literature Review

The complete literature review of the paper deals with 'big data in healthcare' based papers and studies published in scholarly journals focuses on the following objectives:

- Understanding the concept of big data for healthcare.
- Recognizing tools and techniques for big data analytics in healthcare.
- Underlining the future benefits and uses of big data in healthcare.
- Reviewing emerging ethical concerns of big data systems in healthcare.

We obtained most of the pertinent papers used for this study from Research Gate, IEEE, and Google Scholar research sources, which we used to explore for the set of specific articles related to the proposed research. We used an inclusion basis to choose big data and healthcare papers to find relevant papers to answer research questions based on predefined keywords. Our aim is to support developing an emerging ethical framework for Healthcare big data, as shown in Figu

5.2 The Diabetes Dataset (UCL Repository)

Since millions of healthcare data points are created and shared daily, a central data repository that aggregates the entire dataset in one location is needed (Luo et al., 2016). We also need powerful tools to extract information rapidly and analyze the selected data effectively. While Power BI will give healthcare organizations visibility into their data and help them gather many insights, other more effective analytics tools should also be considered. Furthermore, even though the data has been de-identified, there are other ethical issues and concerns that we will discuss in the subsequent section of the article (Durcevic, 2020).

5.3 Research Surveys

In this project, we used primary data collected by the authors' team from randomly selected respondents and a pre-processed dataset originally obtained from Health Information National Trends Survey HINTS 4 Cycles 1 (NCI, 2012).

5.3.1 Primary Research Survey

We used Sogo Survey to conduct the primary research questionnaire to extract respondents' concerns with big data analytics and ensure that the required data is retrievable intelligently.

Unlike the traditional approach, an online survey makes retrieval and analysis of the relevant information more accessible. Power BI visualization is appropriate because it can display complex data in an interactive and user-friendly manner. To move this forward, 53 people filled the survey, and we will address the results in the following segment.

5.3.2 Secondary Research Survey

We used the pre-processed first cycle HINTS 4 survey, conducted on 3959 responders between October 2011 and February 2012, with a response rate of 36.7 percent. Five questions, labeled A-E, were listed, and are discussed further below.

A. Concerns of unauthorized access to their health records as they are transferred electronically between healthcare facilities.
B. Concerned about unauthorized access to their records as they are faxed between healthcare professionals.
C. Satisfied that protections were in place to shield their patient records from unwanted access.
D. Satisfied that they had a voice in collecting, using, and exchanging their medical records.
E. Hidden details from a healthcare provider out of respect for the patient's safety?

We used the following concepts in this work.

6 ANALYSIS AND FINDINGS

6.1 Ethical Problems of Big Data Analytics in Smart Healthcare

As mentioned earlier, this project discusses some emerging ethical concerns of big data in healthcare, including discrimination, data breaching and privacy issues as delineated in the following. We also discuss further how some ethical issues could lead to potential discrimination.

6.1.1 Discrimination

Big data analytics can potentially exacerbate pre-existing demographic gaps in healthcare by presenting biased results from the algorithm used (Cahan et al., 2019; Cutillo et al., 2020). The data used to train these algorithms contributes more to such generalization or stereotypes against a group. Racial biases embedded in typically biased training datasets are more likely to yield racially discriminatory predictive models (Cutillo et al., 2020). For example, the predictive models derived from the Framingham Heart Study and precision medicine protocols centered on European ancestry (Paulus et al., 2018). The causes of discriminatory bias in a dataset could occur at different phases of an analytical pipeline (Suresh and Guttag, 2019).

As observed in Figures 6 & 7 below, the Diabetes readmission dataset used in this project is highly imbalanced. The dataset has an overrepresentation of the Caucasian race, leading to a false generalization. Additionally, there is an aggregation bias in the dataset as it is hard to know which group (race) is others. While the gender feature is well represented, the LGBT populations can feel unfairly aggregated with the two genders.

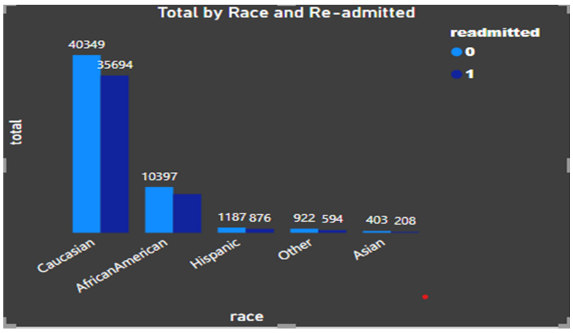

Figure 6: Number of Readmissions by Race.

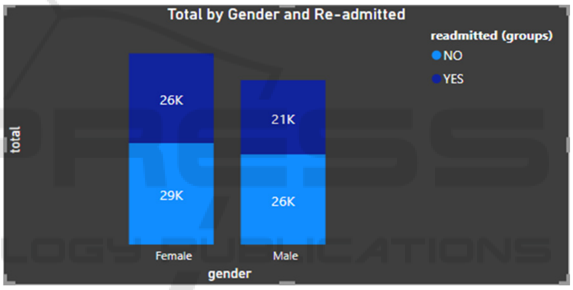

Figure 7: Number of readmissions by Gender.

On the other hand, we cannot say there is an underrepresentation based on age because it is rare for people below 20 years to be diabetic (as shown in Figure 8 below).

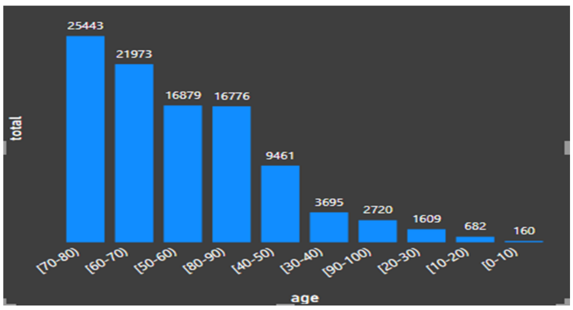

Figure 8: Number of Readmissions by Age.

The model analysis with Power BI Key Influencer (shown in Figures 9 & 10 below) identified evaluation bias. It revealed that Asians, Hispanics, African Americans, and people weighing between 0 and 25kg are unlikely to be readmitted.

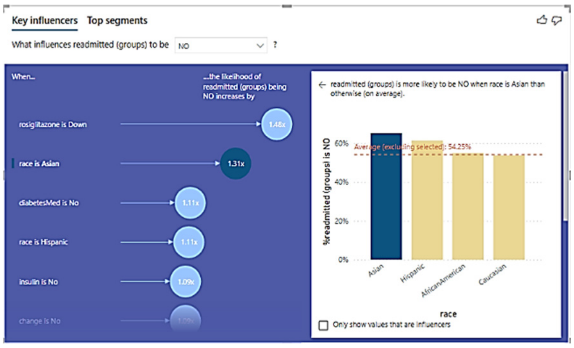

Figure 9: What factors Influences Readmission to be No.

However, Caucasians and people weighing more than 200kg are more likely to be readmitted due to Diabetes. This finding could lead to a misleading generalization of readmitting Caucasian patients weighing more than 200kg even though they are fine.

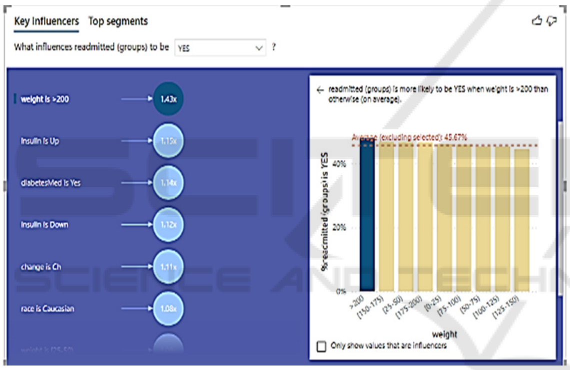

Figure 10: What factors Influences Readmission to be Yes.

Conversely, it might also lead to refusal of admission for patients who do not identify as White or do not weigh up to 200kg.

The aforementioned analytical outcome might lead to social exclusion, marginalization, and stigmatization. Because some persons may be picked out and excluded or included due to the bias, the revelation and application of this study may result in stigma and discrimination. The possible implication could be prioritizing hospital spaces for some patients or refusing to readmit other patients due to their racial identity. The possible implication could also be prioritizing hospital spaces for some patients or refusing to readmit other patients due to their body weight. This finding is consistent with Obermeyer et al.'s (2019) research, identifying how big data analytics could be discriminatory, affecting patients' treatment plans.

6.1.2 Data Breach

Breach of protected health information (PHI) security substantially impacts individuals and healthcare institutions (Agaku et al., 2014). The annual cost of stolen or compromised PHI in the US healthcare sector is estimated to be up to $7 billion. According to research conducted by IBM Security (2019), healthcare data is the most cost of all sectors, with continuous growth in the number of breaches. Figure 11 illustrates that healthcare has the highest average cost of a data breach, almost twice the global average.

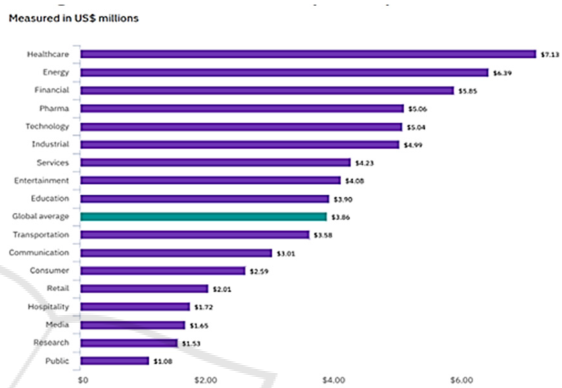

Figure 11: Average cost of data breach by industry.

Big data Ethical challenges are not isolated issues as data breaches could result in the disclosure of personal health information and financial or medical identity theft. In some cases, it can result in severe health consequences on patients (Agaku et al., 2014). Furthermore, a data breach may result in disclosing hitherto undetectable behavioral or psychographic tendencies (Winter, 2018). Data from seemingly insignificant daily routines is gradually being pooled and utilized to uncover behaviors or patterns, clustering or associating individuals into separate groups, resulting in unfair generalizations against such groups. Unauthorized access to private information or activities, such as medical data, could be used to discriminate against persons seeking immigration eligibility, medical treatment, education, banking, and jobs (Winter, 2018).

6.1.3 Privacy

Privacy is a fundamental human right that allows one to choose between exposing or not to expose themselves to others and the rest of the world (Chang, Shi and Zhang, 2019). From the primary survey results shown in Figures 12 & 13, most people agree that big data analytics technologies are functional in healthcare. However, respondents are concerned about the sensitivity of healthcare data, which may jeopardize their privacy.

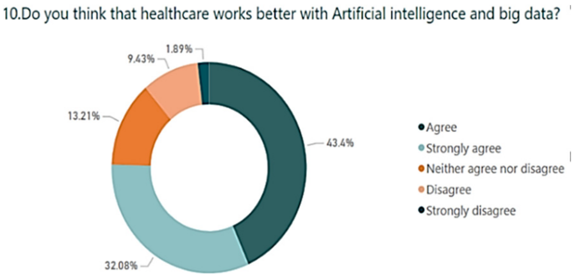

Figure 12: Big data and AI in Healthcare.

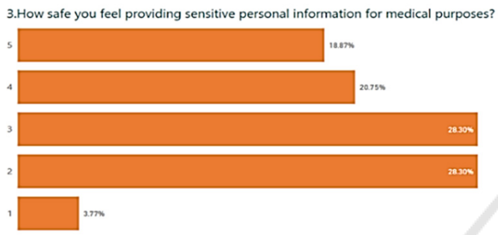

Figure 13: Sensitive Data.

The HINT (NCI, 2020) survey results (shown in Figure 14 below) indicate that, while the majority of respondents are concerned about unauthorized access to their health records, they have confidence that medical providers and institutions would value their voice and therefore keep their data secure.

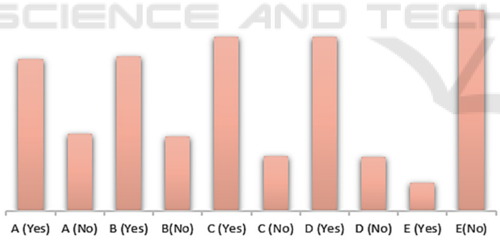

Figure 14: HINT Survey Results.

Personal data can be retrieved at different stages of analytics (Dev Mishra and Beer Singh, 2017). Since modern healthcare services demand patients to provide private and sensitive information to access medical services, clients lose control over the confidentiality of their data when they hand over personal information to third parties and rely on the organization to safeguard its security. Such dependence increases the risk of information leakage if the trusted entity does not implement proper security measures to secure client data (Mariani, Mohammed, and Mohammed, 2015). Protection of patients should be prioritized by avoiding any type of surveillance or unauthorized identification.

7 EVALUATION AND DISCUSSION OF FINDINGS

7.1 Conclusion and Implications

Removing private information to increase patients' anonymity is a powerful method of protecting patient data. The difficulty faced is determining the removable feature with high sensitivity from the data.

While, in cases such as the coronavirus pandemic, the use of sensitive patient data such as location may improve governments' and research institutions' ability to combat the threat more quickly by a surveillance system that provides location data used to curb the current crisis. The diabetes data, on the other hand, has features that could give discriminatory and stereotypical generalizations.

Data scientists must be mindful that utilizing these large amounts of data comes at the expense of human liberty and social autonomy. Lessening the risks of using these data must be monitored by established legislative measures, such as the General Data Protection Regulation (GDPR). The Human-Centered Design approach must be the intent and goals of data usage, including its processing, analysis, warehousing, and dataset sharing.

The following are the main conclusions observed from these principles and criteria for operational use of data-driven healthcare analytics:

7.1.1 Data Sensitivity Is Relative

The description and decision of feature sensitivity vary from project to project, and it also depends on the social value and regulations. For example, the outcome from diabetes data analytics is discriminatory and stereotypical. Can we say that Caucasian white women weighing more than 200 kg are more likely to be diabetic than other ethnic groups?

7.1.2 Discrimination Is Just as Severe

Understanding the significance of data privacy and security is crucial. Most data science ethics journals are concerned with privacy and security and their implications. Notwithstanding, there are discriminatory and racist submissions arising from big data analytics, which also have grave consequences. Furthermore, to ensure a fair model, we must measure analytics discriminatory tendencies against respective advantages.

7.1.3 Human-Centered Design (HCD) Must Be Ethically Compliant

Each phase in the Machine Learning and big data analytics design process should consider the data citizens impacted by models, methods, and algorithms developed by data scientists. Biases in defective datasets, algorithms, and human users are numerous and discussed in depth. We must not ignore that, owing to the vulnerability of data subjects and groups, the risk of discrimination is more severe.

Furthermore, data scientists are also data citizens, asides from developing big data insights, they are also affected by such techniques. As a result, maintaining ethically acceptable data processing and analytics is a win-win scenario for all parties involved.

ACKNOWLEDGEMENT

This work is partly supported by VC Research (VCR0000158) for Prof Chang.

REFERENCES

Agaku, I. T., Adisa, A.O., Ayo-Yusuf, O.A., and Connolly, G.N. (2014) 'Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers', journal of the American Medical Informatics Association, 21(2), pp. 374–378.

Alzahrani, A. G. M., Alenezi, A., Mershed, A., Atlam, H., Mousa, F., and Wills, G. (2020) 'A framework for data sharing between healthcare providers using blockchain'.

Ball, K., Di Domenico, M. L. and Nunan, D. (2016) ‘Big Data Surveillance and the Body-subject’, Body and Society, 22(2), pp. 58–81. doi: 10.1177/1357034X15624973.

Borodin, A. V., Lebedev, N.F., Vasilyev, A., Zavyalova, Y.V., and Korzun, D.G (2016) 'An experimental study of personalized mobile assistance service in healthcare emergency situations', Pdfs.Semanticscholar.Org, (c), pp. 178–183. Available at: https://pdfs.semanticscholar.org/64e7/c0da0b9f7fd403831f6e1a8ade2688f216cc.pdf.

Cahan, E. M., Hernandez-Boussard, T., Thadaney-Israni, S. and Rubin, D.L. (2019) ‘Putting the data before the algorithm in big data addressing personalized healthcare’, npj Digital Medicine, 2(1).

Camilleri, M. A. (2020) ‘The use of data-driven technologies for customer-centric marketing’, International Journal of Big Data Management, 1(1), 50-63.

Cave, A., Kurz, X., and Arlett, P. (2019) 'Real-World Data for Regulatory Decision Making: Challenges and Possible Solutions for Europe', Clinical pharmacology and therapeutics', 106(1), 36–39. https://doi.org/10.1002/cpt.1426

Chang, V., Shi, Y. and Zhang, Y. (2019) ‘The Contemporary Ethical and Privacy Issues of Smart Medical Fields’, International Journal of Strategic Engineering, 2(2), pp. 35–43. doi: 10.4018/ijose.2019070104.

Cutillo, C. M., Sharma, K.R., Foschini, L., Kundu, S., Mackintosh, M. Mandl, K.D., Beck, T., Collier, E., Colvis, C., Gersing, K. Gordon, V., Jensen, R., Shabestari, B. (2020) ‘Machine intelligence in healthcare—perspectives on trustworthiness, explainability, usability, and transparency’, npj Digital Medicine, 3(1), pp. 1–5.

Demirkan, H. (2013) ‘A Smart Healthcare Systems Framework’ in IT Professional, vol. 15, no. 05, pp. 38 - 45, 2013. doi: 10.1109/MITP.2013.35. url: https://doi.ieeecomputersociety.org/10.1109/MITP.2013.35.

Dev Mishra, A. and Beer Singh, Y. (2017) ‘Big data analytics for security and privacy challenges’, Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2016, pp. 50–53. doi: 10.1109/CCAA.2016.7813688.

Dridi, A., Sassi, S.B., Chbeir, R., and Faïz, S. (2020) 'A Flexible Semantic Integration Framework for Fully-integrated EHR based on FHIR Standard' ICAART.

Durcevic, S. (2020) ‘18 Examples Of Big Data Analytics In Healthcare That Can Save People’, Datapine’. Available at: https://www.datapine.com/blog/big-data-examples-in-healthcare/.

European Union Agency for Fundamental Rights (FRA) (2018) ‘Big Data: Discrimination in data-supported decision making’, FRA Focus, p. 14. Available at: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf.

Favaretto, M., De Clercq, E. and Elger, B. S. (2019) ‘Big Data and discrimination: perils, promises and solutions. A systematic review’, Journal of Big Data, 6(1). doi: 10.1186/s40537-019-0177-4.

Floridi, L. (2012) ‘Big data and their epistemological challenge’, Philosophy and Technology, 25(4), pp. 435–437. doi: 10.1007/s13347-012-0093-4.

Francis, J. G. (2014) ‘Introduction: Technology and New Challenges for Privacy,” Journal of Social Philosophy 45(3): 291-303, University of Utah College of Law Research Paper No. 107. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2564314.

Grandview Research Inc., (2021) ‘Healthcare Healthcare Analytics Market Size, Share & Trends Analysis Report,’ Available at https://www.grandviewresearch.com/industry-analysis/healthcare-analytics-market, last accessed on 12/8/2021.

Healthcare Industry Insights (no date) ‘Healthcare Analytics Market Is Estimated To Be Valued At $53’.

Available at: https://sites.google.com/site/healthcare industryinsights/healthcare-medical-analytics-market.
Hermon, R. and Williams, P. (2014) ‘Big data in healthcare: What is it used for?’, Proceedings of the 3rd Australian eHealth Informatics and Security Conference, pp. 40–49. doi: 10.4225/75/57982b9431b48.
Heyman, D. L., and Rodier, G. (2004) 'Global Surveillance, National Surveillance, and SARS. Emerging Infectious Diseases.', PubMed, 10(2), pp. 173–175. doi.org/10.3201/eid1002.031038
IBM (2019) ‘Cost of a data breach report’, IBM Security, p. 76. Available at: https://www.ibm.com/downloads/cas/ZBZLY7KL.
Ioannidis J. P. (2013). 'Informed consent, big data, and the oxymoron of research that is not research', The American journal of bioethics: AJOB, 13(4), 40–42.
Journal, H. (2019) ‘Healthcare Data Breach Statistics’, Www.Hipaajournal.Com, pp. 1–13. Available at: https://www.hipaajournal.com/healthcare-data-breach-statistics/.
Kuempel, A. (2016) ‘The invisible middlemen: A critique and call for reform of the data broker industry’, Northwestern Journal of International Law and Business, 36(1), pp. 207–234.
Luo, J. Wu, M., Gopukumar, D., & Zhao, Y. (2016) ‘Big Data Application in Biomedical Research and Health Care: A Literature Review’, Biomedical Informatics Insights, 8, p. BII.S31559. doi: 10.4137/bii.s31559.
Mariani, D.M.R., Mohammed, S. and Mohammed, S., (2015) ‘Cybersecurity challenges and compliance issues within the us healthcare sector’, International Journal of Business and Social Research, 5(02).
Mittelstadt, B. (2019) ‘AI Ethics – Too principled to fail?’, arXiv, pp. 1–15. doi: 10.2139/ssrn.3391293.
Nair, S. R. (2020). ‘A review on ethical concerns in big data management’, International Journal of Big Data Management, 1(1), 8-25.
NCI (National Cancer Institute) (2020) ‘HINTS 5 cycle 4 public codebook’, Hints. Available at: https://hints.cancer.gov/data/download-data.aspx.
NIH (2021) ‘Health Information National Trends Survey (HINTS)’, National Cancer Institute, p. Survey. Available at: https://hints.cancer.gov/%0Ahttp://hints.cancer.gov/docs/HINTS 2007 Annotated Mail Instrument.pdf.
Obermeyer, Z., Powers B., Vogeli C., Mullainathan S. (2019) ‘Dissecting racial bias in an algorithm used to manage the health of populations’, Science, 366(6464), pp. 447–453. doi: 10.1126/science.aax2342.
Patil, H. K. and Seshadri, R. (2014) ‘Big data security and privacy issues in healthcare’, Proceedings - 2014 IEEE International Congress on Big Data, Big Data Congress 2014, pp. 762–765. doi: 10.1109/BigData.Congress.2014.112.
Paulus, J. K., Wessler, B. S., Lundquist, C. M., & Kent, D. M. (2018) ‘Effects of Race Are Rarely Included in Clinical Prediction Models for Cardiovascular Disease’, Journal of General Internal Medicine, 33(9), pp. 1429–1430. doi: 10.1007/s11606-018-4475-x.
Piai, S. and Claps, M. (2013) ‘Bigger data for better healthcare’, IDC Health Insights, pp.1-24.
Raghupathi, V. and Raghupathi, W. (2020) ‘Healthcare Expenditure and Economic Performance: Insights From the United States Data’, Frontiers in Public Health, 8(May), pp. 1–15. doi: 10.3389/fpubh.2020.00156.
Raja, R., Ali, S., Mukherjee, I., Sarkar, B.K. (2020) ‘A Systematic Review of Healthcare Big Data’, Scientific Programming, 2020. doi: 10.1155/2020/5471849.
Razzak, MI, Imran, M. & Xu, G. (2020) 'Big data analytics for preventive medicine', Neural Comput & Applic 32, 4417–4451. doi.org/10.1007/s00521-019-04095-y
Rehman, A., Naz, S. and Razzak, I. (2021) ‘Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities’, Multimedia Systems. doi: 10.1007/s00530-020-00736-8.
Reinsch, R. W. and Goltz, S. (2016) ‘Big Data: Can the Attempt To Be More Discriminating Be More Discriminatory Instead?’, St. Louis University Law Journal, 61(1), pp. 35–82.
Selby-Bigge, L. A. (ed.) (1975). ‘Enquiries Concerning Human Understanding and Concerning the Principles of Morals’. Oxford University Press.
Sonawane, V. P. and Irabashetti, P. (2015) ‘Method for preventing direct and indirect discrimination in data mining’, Proceedings - 1st International Conference on Computing, Communication, Control and Automation, ICCUBEA 2015, 25(7), pp. 353–357.
Suresh, H. and Guttag, J. V. (2019) ‘A framework for understanding unintended consequences of machine learning’, arXiv preprint arXiv:1901.10002, 2.
Ward, J. S. and Barker, A. (2013) ‘Undefined By Data: A Survey of Big Data Definitions’. Available at: http://arxiv.org/abs/1309.5821.
Winter, J. S. (2018) ‘Introduction to the Special Issue: Digital Inequalities and Discrimination in the Big Data Era’, Journal of Information Policy, 8, 1–4. https://doi.org/10.5325/jinfopoli.8.2018.0001.
Wong, Z. S. Y., Zhou, J. and Zhang, Q. (2019) ‘Artificial Intelligence for infectious disease Big Data Analytics’, Infection, Disease and Health, 24(1), pp. 44–48. doi: 10.1016/j.idh.2018.10.002.
Zhu, H, Wu, C.K., Koo, C.H., Tsang, Y.T., Liu, Y., Chi, H.R. and Tsang, K-F. (2019), 'Smart Healthcare in the Era of Internet-of-Things', IEEE Consumer Electronics Magazine, vol. 8, no. 5, 8822574, pp. 26-30. https://doi.org/10.1109/MCE.2019.2923929
Žliobaitė, I. (2017) ‘Measuring discrimination in algorithmic decision making’, Data Mining and Knowledge Discovery, 31(4), pp. 1060–1089.