

Name: Drishti Harshit Doshi

Student ID: 2203898

Module Code: CS5805

Module Title: Ethics and Governance of
Digital Systems

Academic Year: 2022-2023

Assignment Title: Coursework Assignment

“Individual Analysis and Critical Reflection
on the Use of Digital Technologies in
Healthcare “

Introduction

This assignment study involves analysis and critical reflection on the use of digital technologies in healthcare. The given coursework assignment is from the module CS5805 - Ethics and Governance of Digital Systems led by Dr. Kathy McGrath. The main case study is about the use of AI in radiology work in a hospital setting (Urbanside radiology department). The research papers from the health section provided in Brightspace are reviewed and insights are presented for the same. The ethical and regulatory issues at Urbanside hospital are discussed followed by personal opinions, critical reflection, and analysis of the use of digital technologies in healthcare settings.

Literature Review

A study was conducted to predict pneumonia risk using high-performance generalized additive models (GA2Ms). Multi-task neural nets were contemplated as intelligent tools for identifying symptoms that categorized high-risk pneumonia patients. This helped to predict the probability of death (POD). (Caruana et al., 2015) The models trained on the clinical data learned false assumptions in many instances, for example, the Asthma condition, was recognized as low risk. However, in reality, asthmatics have a high risk of dying. Since asthma patients receive intensive care, their risk of dying is low. These technical blunders can affect the prognosis of the patient. While the deployment of intelligible models yields successful results in the overall prediction of health diagnosis and healthcare facilities, still there are a few crucial setbacks. The missing data in clinical parameters, large error bars, atypical anatomical anomalies, age discrepancies with risk factors, multiple overlapping medical conditions, etc. are some of the causes of inaccurate results. (Caruana et al., 2015) With the use of multiple technologies, intra-technological differences often prove a drawback in achieving end results. Digital technology has been used in terminating the spread of the COVID-19 virus across multiple countries. The Netherlands had major setbacks in using

digital applications for controlling COVID-19 spread. These include weak blending of technologies with the public health system and less engrossment by health professionals. (Sweeney et al., 2021) Further, a study on covid self-regulated tools revealed that the Corona Melder tool was dependent on health workers for intervention, and the Infectieradar tool failed to counsel patients about infection signs. (Sweeney et al., 2021)

The lesser the interdependencies between public health officials and end users, the greater is the effectiveness of COVID-19 transmission prevention. This was noticed in the use of the NHS Covid-19 app by the United Kingdom. This study shares insights for improving the functioning of digital technologies with better user interfaces. The focus should be directed towards "a) advance prescriptive understanding through efficient algorithms, methods, functioning b) advance descriptive understanding through better remark, literature, measurement." (Sweeney et al., 2021) Advancement in an online forum to lower mental stress levels can shield and reduce faith in the use of traditional methods like reinforcement learning, and cognitive behavioral therapy. (Pflügner et al., 2021) One study highlighted two important aspects of user satisfaction for online stress management applications. They include the clinical effectiveness of the application and user-interface features of the application i.e. the kano model. (Pflügner et al., 2021) One of the important issues while designing online stress management applications is dealing with notification alerts. The notification alert feature creates more stress and inhibits the user's focus. How efficient these tools are for reducing stress levels is still a question.

"Teleconsultation", another widespread use of digital services in healthcare refers to "delivering informational facilities to the patient which includes assisting in medical diagnosis, treatment plans, providing emotional and rehabilitative care." (Li et al., 2021) It is interesting to notice how different countries like the United States - Centers for Disease Control and Prevention (CDC), United Kingdom - National Health Service (NHS), China, embraced

teleconsultation services with time. (Li et al., 2021) It was seen that the impact of teleconsultation services had been crucial in delivering patient compliance and it also helped to improve the prognosis during the COVID-19 outbreak. (Li et al., 2021)

Previous studies have shown that the health information exchange lowers the frequency of readmissions, duplicate tests, or medical errors and helps in improving patients' prognoses. (Srivastava et al., 2021) A research study by Oklahoma State University demonstrates "the healthcare referral network model". (Srivastava et al., 2021) The study reveals that the health information exchange forms the underlying base for business value. Health-related agencies are framing business models by using patient-related electronic health data. However, there is a "spillover effect" seen. This is because health professionals are reluctant to share valuable data due to fear of competitive edge.

IBT (Information Boundary Theory) states that the free flow of patients' data is directly proportional to their trust in the AI system and inversely proportional to their privacy concerns. (Yan et al., 2021) With relation to data privacy, PHI (Personal Healthcare Information) disclosure is defined as a collection of generic, general health, and mental health data. (Yan et al., 2021) A study was conducted to address the paradox of privacy and trust with empathy, accountability, and explainability with AI in Depression patients. (Yan et al., 2021) The study revealed that "Empathy" is highly influenced by circumstances and ethnicity across the globe e.g., AI empathy can help to reduce confusion, frustration and increase assurance levels. It is the sole responsibility of the AI designers, developers, and deployers to be "accountable" for the right functioning of the AI systems. (Yan et al., 2021) The government regulatory system should make sure of the appropriate usage of AI tools by defining baseline protocols and standards to ensure an efficient end-user experience. AI empathy can help in the delivery of nurturing and compassionate environment toward a patient. At the same time, technological demands need access to patient-sensitive personal information which can, in turn, affect patient confidentiality and thereby reduce compliance. AI vendor reputation can help to lower the

privacy concerns of the patients. (Yan et al., 2021) Human psychology plays a key role in this. A more responsible, accountable, knowledgeable, honest, and reliable AI company will receive more trust from the people.

"Explainability" refers to knowledge-based trust in the working of the AI system. (Yan et al., 2021) This has a direct influence on patient's acceptance rates of sharing their information. AI technologies have intricate setups which make it difficult to explain their work to the users. However better the explainability, the better the acceptance rate.

Digital technologies like "Fitbit" and "MyFitnessPal" are widely used in today's time. (Hassandoust et al., 2021) Health insurers are using data sets from these smart devices and providing smart services to their customers with business intentions. A study was conducted to analyze the ability of PAYL (Pay-As-You-Live) services among the participants from the United States and New Zealand. (Hassandoust et al., 2021) The study illustrated the framework of the PAYL system. Policyholders' wearable equipment such as fitness bracelets, smartphones, and smartwatches are tracked by the service provider. The health and lifestyle-related data which includes sleep, nutrition, diet, and fitness are shared with the insurance company. If an improvement in the health and lifestyle of the insurer is seen then he will receive benefits like a deduction in insurance premium, reward points, discounts, bonuses, and health counsel. (Hassandoust et al., 2021) This business model has a motivational component. It helps to bring attitude change in people towards a more psychologically balanced life. However, the execution of PAYL services has biased components as well. The service does not consider the socio-economic background of people. Other drawbacks include health data privacy concerns, risk mutualization, means of quantifying individual's health behaviors, user's perspective, interest and compliance. (Hassandoust et al., 2021)

Case Description

Formulating precise diagnoses was an art of high levelled decision-making at Urbanside. (Lebovitz and Sarah, 2019) The efficiency of the report had a

massive impact on both the patient and the referred radiologist. In some situations, the patient's anxiety level (cancer patients) affected the working ability of the radiologist. Indifferences in strengths and weaknesses among radiologists were also an important factor for considering the use of AI technology which would bring equanimity and contribute in eradicating individual diversity in diagnosis. The junior or resident radiologists were less confident in making accurate decisions. Radiologist's name and trustworthiness were at risk with increasing errors in diagnosis. Patient's clinical health records were saved forever with either accurate or inaccurate reports. They served as proof and caused discrepancies among co-workers. Diagnostic errors were broadly categorized into two groups, missing important details and misapprehension of studies. The professionals coined the term "retinol mileage" to portray the confusion created by medical imaging. The interpretation of visual signal translations and their sensitivity formed the core of framing diagnosis. Long-term residency, complex anatomy, and rethinking cases were also contributing factors to diagnostic errors. They also noticed days when radiologists as human beings feel drenched and did not have the full capacity to handle complicated cases increasing the likelihood of errors. The professionals showed acceptance of uncertainty in making unknown mistakes. The extensive mass of data including patient history, and multiple diagnostic films made it difficult to amalgamate editing all the relevant details. Drawbacks of certain radiological imaging techniques also lead to missing intricate human anatomical structures. In cases of life-threatening illnesses, radiologists preferred seeking advice from their colleagues. Fatiguability was an area of concern. Increased fatigue levels were due to time constraints, the large volume of cases, and the diversity in medical imaging data. It caused a decrease in working proficiency leading to more diagnostic fallacies. Urbanside health workers envisioned the strengths of AI tools in improving working efficiency.

The research study was conducted in Radiology Department, at Urbanside Hospital, United States. "Diagnostic radiology is a medical specialty which involves the use of medical imaging technology to help diagnose and treat a

wide range of diseases." In the given study, the data was collected from 34 radiologists from Chest radiology, Paediatric radiology, and Breast imaging. It was observed that one-third of the diagnostic decision involved the use of AI tools. The researcher collected training datasets, codes, and accuracy outputs of the AI model along with medical notes, pictures, and departmental reporting. The grounded theory methods were initially adopted for data analysis followed by a comparison of themes and categories in the available data. Later, data collection and analysis were focused on diseases like "routine breast cancer screening (Breast section), routine lung nodule search (Chest section), and bone age assessment (Paediatrics)." The researcher tried to understand the usage of AI tools by radiologists on basis of their expertise level. There was marked ambiguity seen when "theoretical literature of knowledge production and evaluation processes were studied."

Analysis and Discussion

AI technology could help in boosting the speed of time-bound medical evaluations. The doctors could focus on complex clinical aspects over manual mechanical measuring tasks. The paediatric radiologists at Urbanside hoped to execute an AI device that detects "bone age assessments" which worked on pattern recognitions of patient X-rays using a medical atlas as the reference. (Lebovitz and Sarah, 2019) However, for the majority of the cases, there was a marked variation noticed between the radiologist's provisional diagnosis and AI tool results. The discrepancy in the results prolonged the diagnostic process. Often, it led to disappointment and fatigue in the professionals. They had to repeatedly flip through the literature and machine findings which was indeed a tedious task. The disparity in the findings created confusion and they began doubting the accuracy of their diagnostic results. They started cross-questioning the accuracy, validity, and reliability of AI with their domain knowledge. Urbanside professionals enquired about the AI system's in-depth details like training data used, machine algorithms, and radiologists appointed for training assessments.

Another application of AI included performing routine scans for breast and chest cancers. Radiologists are vulnerable to making mistakes by overlooking finer anatomical details. The scanning process requires a huge amount of expertise as skipping any detail can be life-threatening for the cancer patient. Due to increase ambiguity in the results by CAD devices (computer aided detection tools) and the radiologist's clinical expertise, these tools were used for a second opinion in chest and breast screening. Two noted fallacies that appeared in lung screenings were a) CAD flagged insignificant or least concerned areas and b) CAD skipped areas of consideration or red flags. (Lebovitz and Sarah, 2019) Over time, radiologists began accepting CAD tools working as either false positives or false negatives. They did not question the validity and functioning of the tool. This led to less engagement and enthusiasm among radiologists for exploring the use of AI in health sciences.

The obscurity caused by the use of AI tools in radiology raised suspicion in the professional practice of radiologists. The Urbanside team tried finding effective solutions to deal with AI uncertainty. The paediatric department corrected "AI assessment for 68% cases". The radiologists had multiple opinions on the accuracy of the AI tool which was specific to each case. The senior radiologists often doubted the worthiness of the machine as they had more confidence and faith in their clinical practice for decades. Whereas the junior resident doctors were more dependent on the use of AI tools for the outcomes as they had limited medical knowledge. The results given by AI assessment for bone age were against the gold standard protocols of medical literature. The granular bone age was recorded in the years and the months whereas "Greulich and Pyle atlas" considers only age value. (Lebovitz and Sarah, 2019) Some of the radiologists adapted to the functioning of the AI system. They became open and embraced higher granularity as an appropriate way of finding bone age.

It was a great initiative by Urbanside radiology to embrace and incorporate advanced technology. They had an optimistic vision of helping humanity by

fastening the diagnostic process. However, the goal was not achieved. I do not support implementing testing data in real-time patients because the diagnosis can be manipulated by machine error and this can be life-threatening for the patient. Who would be responsible for that? (Trolley problem) (CS5705 Week 2 Teaching Materials, 2022) The AI tool developers can claim that the technology is to facilitate public service but the impact it has is on the "living human body" and not social media, the economy, or political affairs. This becomes a sensitive ethical issue. It would be much better if the US government took part in the funding. Government initiatives could have advantages like multiple hospitals would be promoted to use AI tools, larger training medical data would be available, the AI tools vendor could collaborate with the Government and work on a larger scale, and the seriousness of the overall thought process of using AI in healthcare (i.e., compliance) would increase.

According to me, the study at Urbanside violates all three principles of utilitarianism, fairness, and human excellence (CS5705 Week 2 Teaching Materials, 2022) because due to the huge outflow of patients and lack of senior staff, the cases were randomly assigned to senior doctors, resident doctors, and junior students. The expertise varies, and hence clinical prognosis of patients is dependent on the doctor who treated them. Also, there were no fairgrounds or set of norms to deal with ambiguous situations caused by AI machines. Each department made its assumptions "universalist ethical stance." (CS5705 Week 2 Teaching Materials, 2022) Some adhered to their professional experience, some relied upon machine reading, and there was overall non-uniformity.

The machine learning models have limited capability to explain how the models work. Ethical issues arising due to machine learning models include transparency, accountability, trust, social inclusion, and safety. Dilemma in the AI Arena is a matter of concern. (CS5705 Week 3 Teaching Materials, 2022) It is declared that accuracy in AI machine-learning models comes at cost of explainability. It is still unclear whether to accept or reject the model when the results are not clarified to users. Does using it violate the right to

explanation legal act? Multinational industries can collaborate to support "Black Box machine learning model" advancements in health. (CS5705 Week 3 Teaching Materials, 2022) Given a scenario where the patient is diagnosed falsely due to machine error, who will take sole responsibility for the patient's critical condition? Legal actions should be taken against medical professionals or AI vendors. The principle of equality is infringed here (CS5705 Week 2 Teaching Materials, 2022) as the innocent patient suffers despite paying the high cost of medical expenses.

Ethical relativism (CS5705 Week 2 Teaching Materials, 2022) applies to Urbanside hospital. The tertiary care unit agreed to work with AI tools because they had pre-set expectations and underlying commitments towards improving the institution's name. The healthcare system is biased, it is not transparent, and easily manipulative. (CS5705 Week 2 Teaching Materials, 2022) There is no higher authority than a doctor to judge the diagnosis of the patient. This can cause a huge number of cases with false diagnoses due to machines, unnoticed or manipulated by doctors. The patient will surely have no clue about the actual truth of his condition. A regulatory system could help to increase the workforce of doctors. Time-related stress and fatigue would reduce, and this will help in better work efficiency and capacity to deal with the ongoing development of the AI system in healthcare.

Reflections

Before a few years the term "Artificial Intelligence" was associated with computational automated tasks whereas, in the current age, there is a shift in the paradigm in using the term AI. AI now refers to the use of technology in "predicting" the outcomes of forthcoming events. (Lebovitz and Sarah, 2019) Researchers have realized massive adversity in the use of AI technology in various domains which includes travel, hospitality, robotics in pharmacy, healthcare settings, online journalism, criminal justice, and radiology. (Lebovitz and Sarah, 2019) In the given research study at Urbanside, the use of AI tools in a hospital setting created a disturbance in

the smooth functioning of the healthcare system. (Lebovitz and Sarah, 2019) The radiologists faced an unexpected, unknown situation when dealing with AI tools. They noticed considerable variation and ambiguity caused by AI. Cross-checking multiple times was time-consuming for the radiologists. As the consequence, their faith in routine practices decreased, multiple queries came up and overall working capabilities decreased. They were attentive towards the working of AI tools and at the same time, not completely reliant on them. (Lebovitz and Sarah, 2019) They could not afford to switch AI tools and medical texts over and over again.

The radiologists struggled to create a balance between the merits of AI and the cost of setting up the AI system. Likewise, medical flaws can be prevalent in teleconsultation services if the assigned work is done by a junior or student medical professional. This can negatively influence the patient engagement ratio. (Li et al., 2021) Concerning people, they need strong back-end support provided by the government regulatory system especially when there are unforeseen situations like abuse or mishandling of PHI (Personal Health Information) by AI vendors. (Yan et al., 2021) There is a paucity in the literature on the understanding connection between "safety, security, fairness, non-discrimination with privacy concerns and trust" as factors of Personal Health Information disclosure. (Yan et al., 2021) A study showed that real-time precise measurement of health-risk prevalence in a given geospatial setting is a challenging task. (Sweeney et al., 2021) There is a need to develop a digital system that overcomes issues of privacy and security for numerous epidemiological risk data. (Sweeney et al., 2021) Another study on online stress management interventions highlighted the development of a high-end user-friendly digital application for patient rehabilitation but at the same time clinical effectiveness is equally important and its significance should not be underrated.

Legislation and policymakers must provide enough socio-economic resources in form of funding and development of facilities that promote the growth of digital systems in health care. They should encourage technological innovations. (CS5705 Week 2 Teaching Materials,

2022) Government systems with the strict formulation of laws and rules for penalizing data offenders will help to subdue patients' privacy concerns. (Yan et al., 2021) Information exchange lies at the heart of the digital transformation (Srivastava et al., 2021) Policymaker companies can initiate business insights with the health industry by using health electronic data of patients. (Srivastava et al., 2021)

Variations in geographical, societal, cultural, behavioral, epidemiological, organizational, psychological, and legal systems across different countries fail to accomplish constancy in the self-regulatory parameters by using digital systems. (Sweeney et al., 2021) Societies having centers of power e.g., press, religions, families, and universities (CS5705 Week 2 Teaching Materials, 2022) must contribute to marketing and raising awareness of the use of AI in healthcare. Crowdfunding, campaigns, and social media platform can be used to promote and collect funds for public welfare.

Let's say the prevalence of AI technology increases in 10 years and all hospital settings use AI systems for diagnosis. In this situation, there is a strong need for digital security to prevent data hacking (Data Protection Act 2018). The U.S. Department of Health & Human Services can bring newer reforms to increase compliance of people towards AI systems in healthcare. Active participation of "Citizens' juries" for dealing with public policy questions. (CS5705 Week 6 Teaching Materials, 2022) This can be implemented by the US Health department for ethical issues e.g., Urbanside radiology AI ambiguity, framing health policies, and advocating community well-being services for patients' mental health.

I would like to progress the research in the MSC Dissertation with the following research topic, "Trust in the deployment of machine learning models in clinical diagnosis: is it a matter of accuracy or explainability. "If the machine learning models are trained and tested on real-world data i.e., real-time patients, then there is surely an underlying risk of uncertainty in the patient's treatment plan. This will affect the patient's prognosis and might prove fatal to their lives. (Caruana et al., 2015) Medical knowledge is vast, uncertain, and incomplete. There are no stable gold standard measures for

accurately training data sets. Understanding domain models and principles are crucial. (CS5705 Week 4 Teaching Materials, 2022) These justifications enhance trust. I would like to study factors that can increase patient compliance and build trust in medical decisions leading to a better prognosis. (CS5705 Week 4 Teaching Materials, 2022)

References

Lebovitz, Sarah, "Diagnostic Doubt and Artificial Intelligence: An Inductive Field Study of Radiology Work" (2019). ICIS 2019 Proceedings.

11. https://aisel.aisnet.org/icis2019/future_of_work/future_work/11, accessed 26 December 2022.

Yan, Aihua and Xu, David, "AI for Depression Treatment: Addressing the Paradox of Privacy and Trust with Empathy, Accountability, and Explainability" (2021). ICIS 2021 Proceedings. 15.

https://aisel.aisnet.org/icis2021/is_health/is_health/15

Srivastava, Ankita; Bao, Chenzhang; and Delen, Dursun, "Assessing Ambulatory HIT Spillover Effects on Hospital Inpatient Costs" (2021). ICIS 2021 Proceedings. 8. https://aisel.aisnet.org/icis2021/is_health/is_health/8

Sweeney, Jeffrey and Vervest, Peter, "Disrupting Viral Transmission by Facilitating Guided Self-Regulation" (2021). ICIS 2021 Proceedings. 13.

https://aisel.aisnet.org/icis2021/is_health/is_health/13

Li, Yiqing and Tong, Yu, "Exploring the Effects of Leader's Actions on Patients' Continuous Engagement in Team-based Teleconsultation" (2021). ICIS 2021 Proceedings.

11. https://aisel.aisnet.org/icis2021/is_health/is_health/11

Caruana, R., Lou, Y., Gehrke, J., Koch, P., Sturm, M. & Elhadad, N. 2015, "Intelligible Models for HealthCare: Predicting Pneumonia Risk and Hospital 30-day Readmission", ACM, , pp. 1721

Pflügner, Katharina; Maier, Christian; Hielscher, Maren; and Weitzel, Tim, "Online Stress Management Interventions: The Role of Application Features" (2021). ICIS 2021 Proceedings.

12. https://aisel.aisnet.org/icis2021/is_health/is_health/12

Hassandoust, Farkhondeh; Johnston, Allen; and Singh, Tripti, "Smart Pay-As-You-Live Services in Healthcare: A Balance Theory Perspective" (2021). ICIS 2021 Proceedings. 6. https://aisel.aisnet.org/icis2021/is_health/is_health/6

CS5705 Week 2 Teaching Materials (2022). Lecture 2: Introducing the ethical, governance and regulatory contexts. Available at <https://brightspace.brunel.ac.uk>, accessed 22 December 2022.

CS5705 Week 3 Teaching Materials (2022). Lecture 3: Accuracy versus explainability in AI. Available at <https://brightspace.brunel.ac.uk>, accessed 27 December 2022.

CS5705 Week 4 Teaching Materials (2022). Lecture 4: Substantive Ethical Issues in Contemporary Contexts. Available at <https://brightspace.brunel.ac.uk>, accessed 26 December 2022.

CS5705 Week 6 Teaching Materials (2022). Lecture 6: Engaging the public. Available at <https://brightspace.brunel.ac.uk>, accessed 28 December 2022.

