Article IV

Salami, Emmanuel (2019) AI, Big Data and The Protection of Personal Data in Medical Practice

Reprinted from European Pharmaceutical Law Review, Vol 3 Issue 4 (2019) 165 – 175 with the permission of Lexxion Publishers.

AI, Big Data and The Protection of Personal Data in Medical Practice

*Emmanuel Salami**

Medicine as an area of endeavor is very important to humanity because of its significance to the healthy preservation of human life. Overtime, various devices with varying levels of sophistication, (dependent largely on the era), have been developed with the aim of providing support to medical professionals in the delivery of medical services.[1] *As of today, Artificial Intelligence (AI) has also lent its support to the development of medical devices by not only making such devices smarter,*[2] *but in some cases being the device itself. As will be discussed in later parts of this article, so sophisticated is AI that it could even assist in the performance of medical surgeries.*[3] *Two important components of AI's role in medical practice are natural language processing*[4] *and computer vision*[5] *which are play key roles in patient assessment and diagnosis.*

I. Introduction

The introduction and use of AI and big data is fast becoming a norm across various sectors of the global economy and the health sector is no exception. As of today, AI is being deployed for disease diagnosis, clinical trial research, drug discovery, medical consultations to mention just a few. The implication of the usage of AI in this manner is that large amounts of personal data will be processed by AI systems designed for use in the medical space. This Article seeks to examine the various uses of AI in medical practice within the European Union/European Economic Area with the objective being the analysis of potential privacy challenges that may arise therefrom.

The concepts of AI and big data are interwoven and very fundamental to the topic at hand. AI uses large volumes of big data in order to (among other things), make AI learn from its experiences, a process known as machine learning.[6] The implication of this is that large volumes of big data are collected by relevant AI systems with quite a number of data protection consequences. From a medical perspective, the possible consequences of AI's big data processing are further heightened by the sensitive nature of the medical data processed therein.[7]

In order to address some of these concerns, this article seeks to discuss the possible challenges that could arise from the use of AI and big data in the practice of medicine. To achieve this, the basic concepts of AI and big data will be briefly discussed with examples being made of how AI is being used in the practice of medicine. The (potential) privacy concerns in the use of AI and big data in medical practice will be considered with possible recommendations that may resolve these concerns being made.

DOI: 10.21552/eplr/2019/4/7

* Emmanuel Salami, Doctoral Candidate in Information Technology/Data Protection Law at the Faculty of Law, University of Lapland, Finland. For correspondence: <Esalami@ulapland.fi> The author would like to thank Professor Rosa Maria Ballardini and Professor Rob van den Hoven van Genderen for their useful comments and review of this article.

1 World Health Organization, *Medical Devices: Managing the Mismatch: An Outcome of the Priority Medical Devices Project* (WHO Press Switzerland 2010).

2 This may be through the use of (medical) Internet of Things (IoT). IoT has been defined as a network of physical devices and other items, embedded with electronics, software, sensors, and network connectivity, which enables these objects to collect and exchange data. For further reading, see, Dimiter Dimitrov, 'Medical Internet of Things and Big Data in Healthcare' (2016) 22 Healthc Inform Res 3, 156 - 163.

3 ibid 11.

4 Jerry Kaplan, *Artificial Intelligence - What Everyone Needs To Know* (Oxford University Press 2016) 60 - 64.

5 ibid 54 - 57.

6 Machine learning is the scientific study of algorithms and statistical models used by computer systems to perform specific tasks through reliance on patterns and inferences. For further readings, see, Peter Flach, *Machine Learning: The Art and Science of Algorithms that Make Sense of Data* (*Cambridge University Press* 2012).

7 The definition of sensitive personal data can be gleaned from the provision of Art 9 GDPR which lists personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation etc. as special categories of data.

The main goal of this research will be to ensure the protection of personal data in the use of AI and big data in medical practice. The definition of AI, big data and their usage in medical practice are considered subsequently.

II. When Does a Computer Become intelligent?

Alan Turing defined intelligent behavior (in computers) as the ability to achieve human-level performance in all cognitive tasks, sufficient to fool an interrogator. For Turing, the key test of computer intelligence is when the results generated by the computer cannot be distinguished from those of its human counterparts.[8] Russel and Norvig also define AI as 'the study of agents that exist in an environment and perceive and act'.[9] These definitions reflect the fact that AI has been developed to largely imitate human beings in the way they act and make decisions, thereby possessing an intelligence which in itself is modelled after human intelligence. An important tool in the intelligence of computers is big data. The relationship between AI and big data stems from the fact that AI uses big data sets to learn from its experiences (machine learning) so that AI can carry out repetitive tasks based on its past activities. Big data has been defined in terms of 'three Vs' where 'Volume' relates to massive datasets, 'Velocity' relates to real-time data and 'Variety' relates to different sources of data.[10] Though there is no one size fits all definition of big data, it is safe to describe big data as data which is difficult to process using traditional processing methods.

Having examined the above definitions, some case studies of AI being developed and used in the medical space are considered as follows:

1. Disease Diagnosis and Prediction

One of the most noticeable strides being made by AI is in the area of disease diagnosis and prediction. As of today, AI is capable of classifying echocardiograms[11] in order to determine whether they reveal any known heart conditions. When it comes to classifying echocardiograms, AI outscores human beings in terms of the accuracy of its classifications at a percentage of 92% as against the 79% accuracy of human beings.[12] Beyond classification, AI is also capable of predicting which patients are more likely to have a heart attack within a ten year period by scanning the patients' routine medical data with established risk factors such as high blood pressure, cholesterol, smoking, age etc. As of today, there exists an AI system which has correctly predicted 355 cases of heart attacks.[13]

2. Telerobotic Surgeries and Robot-Assisted Surgeries

Telerobotics, considered to be an integral part of the wider field of telemedicine,[14] aims to provide specialized healthcare services over long distances, effectively eliminating the need for the physical presence of both the physician and the patient at the same location.[15] These surgeries are carried out with the aid of robotically controlled instruments through which coronary intervention can be undertaken by a doctor without any physical contact with the patient.[16] AI

8 Alan Turing, 'Computing Machinery and Intelligence' (1950) 433 - 460 <https://www.csee.umbc.edu/courses/471/papers/turing.pdf > accessed 24 August 2019.

9 Peter Norvig und Stuart Jonathan Russell, *Artificial Intelligence: A Modern Approach* (Pearson 2010) 7.

10 Doug Laney, '3-D Data Management: Controlling Data Volume, Velocity and variety' <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> accessed 20 October 2019.

11 Echocardiograms are used to paint the picture of a patient's heart with the picture being subsequently interpreted by cardiologists so as to identify any medical conditions. For further readings, please see, Michael Quartermain, 'Echocardiogram' in Ernerio Alboliras et al (eds), *Visual Guide to Neonatal Cardiology* (Wiley Online Library 2018).

12 Ali Madani et al, 'Fast and Accurate View Classification of Echocardiograms Using Deep Learning' (2018) 1 Digital Science 6.

13 Stephen Weng et al, 'Can Machine Learning Improve Cardiovascular Risk Prediction Using Routine Clinical Data?' (2017) PLOS.

14 For further readings on telemedicine, please see, Bernard Dickens and Rebecca Cook, 'Legal and Ethical Issues in Telemedicine and Robotics' (2006) 94 International Journal of Gynecology & Obstetrics 1, 73-78.

15 Sotiris Avgousti et al, 'Medical Telerobotic Systems: Current Status and Future Trends' (2016) 15 BioMedical Engineering Online 96.

16 The first telerobotic surgery was performed in New York by a team of French doctors on a patient in Strasbourg in the year 2001. The telerobotic surgery dubbed 'Operation Lindbergh'. A similar surgery was performed in the year 2018 in India from a distance of 32 km. See, Mahesh Langa, 'Ahmedabad Doctor Performs Telerobotic Surgery on Patient 32 km Away' *The Hindu* (Chennai December 2016).

is also being developed to assist in the delivery of medicare by monitoring the body organs (like the heart, lungs etc.) and providing first response to patients in hospitals where possible.[17]

3. Genomics

Genomics is the study of all of a person's genes (the genome), including interactions of those genes with each other and with the person's environment.[18] Gene editing (a subset of genomics), is the ability to make specific changes in the DNA sequence of a living organism (for the purpose of this article, a natural person) in order to remove certain traits therefrom and insert new traits.[19] Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR-Cas9) is a key molecular tool that serves as a gene editing technology.[20] As of today, AI systems have been developed for carrying out gene editing in a cost-effective manner.[21]

4. Drug Discovery, Development and Repurposing

AI is being used for drug design a practice which now eases the drug discovery process. Traditionally, the first step in drug development is understanding the biological origin of the relevant disease as well as its resistance mechanisms.[22] This is followed by the identification of proteins that could help in curing the disease. AI is now capable of analyzing all available data as well as the identification of good target proteins.[23] Drug repurposing, as the name implies, is the application of an approved drug for the treatment of a different disease. By evaluating the available data of drug molecules to match new targets, AI now plays an integral role in drug repurposing.[24] AI systems that are used for this service contain billions of documented interactions among patients, existing drug molecules and rare diseases.[25]

5. Clinical Trials

To put it simply, clinical trials are investigative medical research studies where the potency, side effects etc. of a new drug are tested on people.[26] One of the hardest stages of the clinical trial process is the choosing of right candidates to participate in such trials. So important is this process that selecting the wrong candidates may prolong clinical trials at huge monetary costs. AI is now capable of identifying suitable candidates for specific clinical trials to the immense benefit of the clinical trial process.[27]

Flowing from the decision of the Court of Justice of the European Union (CJEU) in case C-329/16,[28] the CJEU had interpreted the definition of medical devices, (under Article 1(1) and Article 1(2)(a) of the medical device directive,[29] to include self-standing software.[30] In accordance with the decision of the

17 Behnood Gholami et al, 'AI Could Provide Moment-by-Moment Nursing for a Hospital's Sickest Patients' <https://spectrum.ieee.org/biomedical/devices/ai-could-provide-momentbymoment-nursing-for-a-hospitals-sickest-patients> accessed 20 October 2019.

18 National Human Genome Institute, 'A Brief Guide to Genomics' (2015) <https://www.genome.gov/about-genomics/fact-sheets/A-Brief-Guide-to-Genomics> accessed 26/10/2019.

19 Judith Fridovich-Keil, 'Gene Editing' (Encyclopedia Britanicca 2019) <https://www.britannica.com/science/gene-editing> accessed 26 October 2019.

20 ibid.

21 For further reading on AI being developed for gene editing, see Jade Sterling, 'A Step Closer to Gene Editing with AI' (*Khalifa* 2019) <https://www.ku.ac.ae/a-step-closer-to-genome-editing-with-ai/> accessed 26 October 2019.

22 Lothar Terfloth et al 'Drug Discovery: An Overview' in Thomas Engel and Johann Gasteiger (eds), *Applied Chemoinformatics* (Wiley 2018).

23 Rakesh Sharma, 'AI- The Future of Pharma Industry' (*ExpressPharma* 2019) <https://www.expresspharma.in/it-at-pharma/artificial-intelligence-the-future-of-pharma-industry/> accessed 20 October 2019

24 ibid.

25 ibid.

26 Melissa Conrad Stöppler, 'Clinical Research and Clinical Trials' (*MedicineNet* 2019) <https://www.medicinenet.com/clinical_trials/article.htm#clinical_research_and_clinical_trials_facts> accessed 29 October 2019.

27 Peter Meath, '3 Ways AI Could Transform Clinical Trials' (*JP Morgan* 2018) <https://www.jpmorgan.com/commercial-banking/insights/ai-transform-clinical-trials> accessed 20 October 2019.

28 Case C-329/16 *Syndicat national de l'industrie des technologies médicales (Snitem) & Philips France vs Premier Ministre & Ministre des Affaires sociales et de la Santé* [2017] EUECJ C-329/16.

29 Council Directive 93/42/EEC of 14 June 1993 concerning medical devices, as amended by Directive 2007/47/EC of the European Parliament and of the Council of 5 September 2007. See also: Case C-219/15 *Elisabeth Schmitt vs TÜV Rheinland LGA Products GmbH* [2017] ECLI:EU:C:2017:128.

30 Case C-329/16, para 40.

CJEU, this definition of medical devices remains the same even when it does not act directly in or on the human body.

Furthermore, for these AI systems to be 'trained' to carry out specific tasks, they have to pass through the machine learning phase which requires the uploading of large volumes of (sensitive) personal data into AI systems. There are also large volumes of personal data transmission both online and offline, uploading of personal data on various computer systems and IoT to mention just a few. Activities of this nature pose serious data protection concerns for reasons which include the exposure of sensitive personal data to the attendant risks of AI and big data processing operations. It is therefore important to address these concerns before they result in a breach of the right to data protection of data subjects.[31]

III. Data Protection Concerns in the Use of AI and Big Data in Medical Practice

The processing of sensitive personal data by AI systems generally poses substantial concerns in most sectors of the economy and the medical space is no exception. As indicated above large personal data sets are required not just to implement the machine learning phase of AI, but also for the daily operation of the AI systems. Some visible challenges posed by the use of AI and big data in medical practice are highlighted below:

1. The Lawfulness and Transparency Dilemma

The requirements that personal data should be processed lawfully and transparently are fundamental principles of data protection law.[32] These principles respectively require that personal data should be processed pursuant to a justifiable legal basis[33] and that the nature and purpose of the processing activity should be clearly communicated to the data subjects. The basic rule as it pertains to the lawfulness and transparency principle is that the purpose of the processing activity must have been identified and defined at the commencement of the processing activity as it will form part of the information[34] that will be provided by the Controller[35] to the data subject.[36] From a big data perspective, these requirements may not be so easy to meet. This is partly because the determination of an appropriate and justifiable legal basis for processing personal data may raise some concerns such as the necessity of even processing such categories of health data within the scope of data protection law, eventually making it unnecessary for any legal basis to be put in place. A good illustration of this challenge can be seen in the processing of genetic data by AI systems. In such processing activities, researchers have found that the possibility of identifying data subjects from 'anonymized' genetic data exists despite such 'anonymization'.[37] In such a case, it is therefore necessary to determine if the personal data in question falls within the purview of data protection law as well as the appropriateness of the legal basis for processing the personal data in the event that data protection law is applicable. Another concern is that, as previously mentioned, AI processes large volumes of big data with a high tendency to discover further purposes for data processing (known as the repurposing of data) which were not envisaged or existent at the beginning of the processing activity.[38] In the repurpos-

31 The data subject is the person whose personal data is being processed during a processing activity. In the context of this article, data subjects would be patients or persons subject to the medical examination, decision making and/or surgical operation of medical AI systems.

32 Art 5 (1) (a) GDPR. See also: Lee Bygrave: '*Data Privacy Law, an International Perspective*' (Oxford University Press 2014) 176.

33 Arts 6 and 9 of the GDPR for personal data and sensitive personal data respectively.

34 In accordance with the transparency requirement, Art 13 of the GDPR requires that data subjects ought to be adequately informed about the nature and purpose of the processing operation at the point of personal data collection.

35 Art 4(7) GDPR defines a controller as the person who determines the means and purpose of the processing activity. In simple terms, this is the person who determines the 'what', 'why', 'who', 'where' and 'how' of a processing activity.

36 Art 29 Data Protection Working Party, Guidelines on consent under Regulation 2016/679, 17/EN WP259 rev.01, 23, para 6.

37 Mark Phillips, 'Can Genomic Data be Anonymized?' (*Global Alliance* 10 October 2018) <https://www.ga4gh.org/news/can-genomic-data-be-anonymised/> accessed 5 December 2019; See also, Megan Molteni, Genome Hackers Show No One's DNA is Anonymous Anymore (*Wired* 10 November 2018) <https://www.wired.com/story/genome-hackers-show-no-ones-dna-is-anonymous-anymore/> accessed 5 December 2019.

38 For further reading, see, Robin Pierce, 'Machine Learning for Diagnosis and Treatment: Gymnastics for the GDPR' (2018) 4 EDPL 3, 339 – 340; Mary Shacklett, 'Repurpose Big Data to Get More Analytics Bang for Your Bucks' (*Tech Republic* 28 January 2014) <https://www.techrepublic.com/article/repurpose-big-data-to-get-more-analytics-bang-for-your-bucks/> accessed 31 October 2019.

ing of big data, the processing of personal data on a justifiable legal basis as well as informing data subjects about the new purposes for the processing operation may pose a challenge of some sort for data controllers. This is because the initial legal basis that was used for the processing activity may not be applicable for the further use that is uncovered through the machine learning process. In data repurposing, adequate information about the nature and purpose of the processing activity must also be clearly provided in a manner that the data subject must know when to expect privacy and when he or she may part with his/her privacy.[39]

2. Decision Making Algorithms

Another data protection concern that arises in the use of AI in medical practice is the automated decision making capability of AI systems 'which produces legal effects that affect natural persons'.[40] In its use in medical practice, AI will make quite a number of decisions which will significantly have legal effects on data subjects. For instance, in disease diagnosis, AI could potentially be the basis for deciding that a person is not physically or medically fit to work in a particular position. In order to be compliant with data protection law, data subjects ought to be provided with adequate information[41] about the processing activity and also the logic deployed by the AI system in reaching its decision ought to be provided to the data subjects whose legal rights may be affected by such decisions.[42] There is also the need to have such decision reviewed by a natural person so as to consider the correctness of the decision and potentially correct any errors therein.[43] The importance of these legal requirements are further underlined by the fact that AI is capable of making inferences[44] from big data through big data analytics,[45] with such inferences having the potential to affect the legal rights and freedoms of natural persons. For instance, in the case of an AI system which predicts that persons may have a heart attack in about 10 years, inferences may be drawn from this prediction which may identify certain factors (such as medical history, race, living conditions etc.) which are predominant among such persons. These inferences may be used in future to determine the legal rights of data subjects and in the failure to provide the aforementioned statutory information, explanation (when requested) or human intervention, such decision-making process will be in infraction of data protection law as AI systems will play a critical role in determining the fate of data subjects without the data subjects having an understanding of how the AI system reached its decision or even in some cases, data subjects may have no idea that AI systems are capable of reaching such decisions. Unfortunately, not all AI systems are capable of complying with the above legal requirements either due to the reluctance of data controllers to comply or in some other cases, the non-inclusion of the principle of privacy by design[46]

39 The principle of the 'expectation of privacy' is a principle of privacy law that (among other things), determines in which places and in which activities a person has a legal right to privacy. This principle implies that the unreasonable compromising of another's interest in keeping his/her affairs from being known can make the intruding party liable for the breach of the other's privacy. This principle has been addressed by the European Court of Human Rights in a plethora of cases which includes the case of Copland vs United Kingdom, Application no. 62617/00, ECHR, 3 Apr 2007.

40 Art 22 (1) GDPR; The decisions being made by AI systems are derived from their algorithms which have been trained through machine learning to make specific decisions under identified circumstances. Algorithms can be described as a set of rules or instructions that takes input data and converts it to output data. For further reading, see, Claude Castelluccia and Daniel Le Métayer, 'Understanding Algorithmic Decision-Making: Opportunities and Challenges' (Brussels, Panel for the Future of Science and Technology 2019) <http://www.europarl.europa.eu/RegData/etudes/STUD/2019/624261/EPRS_STU(2019)624261_EN.pdf> accessed 31 October 2019.

41 Art 13 GDPR.

42 ibid; It is worth mentioning that the existence of the right to an explanation of automated decisions has been questioned with doubts raised as to its existence under the GDPR. See, Sandra Wachter, Brent Mittelstadt, and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 2, 76 - 99.

43 Recital 71 and Art 22 GDPR.

44 Inferences have been defined 'as information relating to an identified or identifiable natural person created through deduction or reasoning rather than mere observation or collection from the data subject'. For further readings, see, Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2018) Columbia Business Law Review.

45 Big data analytics involves examining large volumes of big data so as to uncover specific information -such as hidden patterns, correlations, market trends and customer preferences - that can help organisations make informed business decisions; Michael Minelli et al (2013) 'Business Analytics' in Michael Minelli et al (eds), *Big Data, Big Analytics* (Wiley 2018).

46 The principle of privacy by design requires that data protection considerations should be included in a processing activity when the means for such processing activity are being determined. For further readings, see, COE/FRA: Handbook on European Data Protection Law (Luxembourg, Office of the European Union, 2018) 183-184.

which results in the development of algorithms which are incapable of meeting the legal requirements.[47]

3. Data Security and Data Access

Due to the large volume of big data being processed by AI systems coupled with the sensitivity of such personal data, its security is a very fundamental topic. The security of personal data being used by AI systems in the medical space could be further constrained by the fact that personal data is needed in some cases in test environments. This is because in order to carry out the machine learning process through which AI systems are trained to perform their tasks, large volumes of big (personal) data are uploaded into AI systems and except there are proper security measures in place, this may impact negatively on the confidentiality, integrity, availability and resilience of processing systems and services. There is also the tendency that personal data could potentially be accessible to more people than would ordinarily have access to it in the traditional practice of medicine. For instance, in Telerobotic surgeries and robot- assisted surgeries, engineers and information security experts will be some of the experts that will be needed as part of the team for such surgery due to the fundamental nature of computer systems and technology to that process. Sensitive personal data will also be uploaded on different computer systems and will be transmitted both online and offline. This could raise serious concerns like the hacking of such systems, unauthorized access amongst other non-medical personnel experts on (or even outside) the team who should not have access thereby granting data access to unauthorized persons in infraction of the extant data protection laws.[48]

4. Discrimination of Data Subjects

Another challenge that arises from the use of AI in the medical space is the discrimination that data subjects could potentially suffer based on the biases that have been inputted into the algorithms of the AI system during the machine learning phase. Article 14(1) of the European Convention on Human Rights (ECHR) guarantees the enjoyment of rights and freedoms set forth in the ECHR without any discrimination on the grounds of sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status. The fact is that any AI system that has been designed to classify, rate or produce any useful result to justify any decision is bound to discriminate in the sense of making distinctions between people based on certain features. However, as reflected in Article 14 ECHR, there are some grounds of discrimination which are prohibited under the law such as race, ethnic origin, religion, political beliefs, gender, sexual preference etc.[49] There are yet to be standardized and generally accepted thresholds regulating the development of algorithms. The implication of this is that these algorithms could therefore be subject to the biases of the engineers, system, processes, non-divergent data categories and even non-divergence in the data subjects that are used at the development and trial stage with varying consequences of grave proportions for unrepresented or underrepresented data subjects.[50] A real-life example of such algorithmic discrimination occurred in the USA where an algorithm that was used to allocate healthcare to patients was systematically discriminating against people of colour by allocating lower risk scores to them even when they suffered the same ailments as their white counterparts. As a result, such people of colour were less likely to be referred to programmes that provided more personalized care.[51] Such discriminatory occurrences have

47 Three approaches in the design of algorithms have been identified as being critical in the provision of explanations to data subjects. The 'black box' approach is the first of the three approaches and under this approach, the algorithm can be analyzed through an observation of the input and output of its system without any knowledge of its code. The second approach in the design of algorithms is known as the 'white box' approach under which algorithmic decisions can be explained using its code. The third approach known as the 'constructive approach' is the approach which takes the need to provide explanations into the consideration when AI systems are being designed.

48 Art 5 (1) (f) GDPR and Art 32 GDPR.

49 See also, Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation; Convention for the Protection of Human Rights and Fundamental Freedoms as amended by Protocols No 11 and No 14.

50 Castelluccia and Le Métayer, 16 and 17.

51 Ziad Obermeyer et al, Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations (2017) Science 336, 447-453-

the potential to deprive data subjects of core EU rights such as the respect for human dignity, freedom, equality, the rule of law and respect for human rights to mention just a few.[52] If this discrimination that could be occasioned by AI goes unchecked, it has the potential to wipe out the benefits of AI in one breath as the courts are usually unhesitant in nullifying any form of discrimination.[53]

5. Data Transfers

There is the possibility of data transfers occurring in the use of AI in the medical space. One of the most interesting scenarios where this is possible is in the use of telerobotic surgeries. As described above, telerobotic surgeries involve the execution of surgeries without the surgeon and the patient being in the same location with the surgery being executed through the use of robotically controlled instruments. The implication of this is that the personal data of the patient will be transferred on robots and computer systems which may amount to the transfer of personal data in some cases. For instance, the first telerobotic surgery 'Operation Lindbergh' was carried out on a patient in France while the medical team was located in the USA.[54] In a situation where telerebotic surgeries are sought to be carried out in a jurisdiction outside the EU/EEA, particularly in countries without an adequacy decision or privacy shield,[55] while patients remain in the EU/EEA, uploading personal data on computer systems outside the EU/EEA will become inevitable, thus amounting to data transfers. In such a case, except the specific laws of the European Union as it concerns data transfers are complied with, there is bound to be a compliance gap arising therefrom.

IV. How do we Resolve These Concerns?

Having identified the concerns raised by the application of AI and big data in the medical space, it is necessary to seek effective measures to tackle the identified concerns.

1. The Lawfulness and Transparency Dilemma

In respect of the concerns raised by the lawfulness principle, data controllers must ensure that either before commencing a processing activity or at the point where new purposes are being identified therein, that a justifiable legal basis is put in place. The conditions listed in Article 6 (4) GDPR which relate to the repurposing of data ought to be strictly adhered to whenever further purposes for processing big data are identified. The categories of personal data in question should also be examined and the possibility of having such data categories falling within the purview of data protection law considered. For instance, in the case of genetic data, despite its 'anonymization', such data categories must also be examined on a case by case basis so as to, (with more accuracy), determine whether natural persons can be identified therefrom. It is only when it is certain that natural persons cannot be identified that the requirement for a justifiable legal basis can be dispensed with, otherwise, controllers should ensure that a lawful and justifiable legal basis is provided for all data categories. In order to comply with the transparency principle, data subjects must be adequately informed about the nature and the purpose of the processing activity. For instance, if there is a repurposing of data and the purpose of the processing activity is changed into some other purpose which was not anticipated at the beginning of the processing activity, then the provision of Article 6 (4) GDPR ought to be applied in order to determine if the processing activity can be processed under the existing legal basis or if a new legal basis will be required along with the provision of additional information about the new purpose (if necessary). In this manner, the 'expectation of privacy' requirement would have been met

52 These are some of the rights enshrined in the European Charter on Human Rights and other similar treaties that are applicable in the EU/EEA.

53 For an elaborate discussion of the cases of the International Court of Justice, the CJEU, European Court of Human Rights etc. on discrimination, see: COE/FRA: Handbook on European non-discrimination law, (Luxembourg, Office of the European Union, 2018) 183-184.

54 (n 16).

55 The Privacy Shield is a legal framework designed to safeguard the transfer of data from the EU and Switzerland to the USA. For further readings, see, 'An Overview of the Privacy Shield' <https://www.privacyshield.gov/Program-Overview> accessed 29 October 2019.

and data subjects will be able to anticipate every potential effect of the processing activity.[56]

2. Decision Making Algorithms

To be compliant with the extant law,[57] the decision-making process of algorithms ought to be capable of being explained so that data subjects who are subject to such decisions and whose legal rights will be affected in the process may through such an explanation understand the underlying logic behind the decision that has been reached against them. Unfortunately, this may not always be the case as not all algorithms may be capable of having their decisions explained.[58] In order to resolve this dilemma, it is necessary to, (with the aid of the principle of privacy by design), integrate into the codes of AI algorithms, sufficient means and processes which permit an explanation of the decision making process of AI systems that will be used in the practice of medicine. This can be achieved through the adoption of the 'constructive approach to the design of algorithms' which takes algorithmic accountability into consideration during the design phase of algorithms. There is also a need for the standardization of the development of algorithms that control the decision-making process of AI systems so as to ensure the adoption of practices and processes which support and encourage algorithmic accountability. Another important component necessary to achieve compliance with the requirements of data protection law as far as the decision making process of algorithms is concerned is the requirement that the decisions of AI systems should be subject to human review.[59] This is to ensure that natural persons do not have their fate subjected solely to the decision making capabilities of an AI system thereby giving up their autonomy in the process.[60] Therefore, some mechanisms which will ensure that medical decisions made by AI systems are subject to one form of human intervention or oversight ought to be put in place. To amount to sufficient human intervention or oversight, controllers must ensure that persons exercising the human intervention have the knowledge, authority and competence to change decisions that have been reached by AI systems.[61] These recommendations can be effectively implemented and monitored through the standardization and strict regulation of the use of AI in the medical space through (for instance) the setting up of a regulatory body as will be discussed subsequently.

3. Data Security and Data Access

The GDPR does not expressly list the processes to be adopted in order to ensure the security of personal data. Rather, it lists the objectives which ought to be achieved through the adoption of adequate technical and organizational measures.[62] However, due to the large volume and sensitivity of the big data being processed under this processing activity, it is necessary to have the technical and organizational measures standardized possibly through the adoption of uniform measures for specific processing activities so as to ensure the protection of personal data within the medical space. This can be achieved through the creation of a certification body that will be responsible for the regulation of AI systems designed for use in the medical space. Such a body will design the necessary safeguards which subject to the relevant provisions of the GDPR, will sufficiently guarantee the security and access rights to the sensitive personal data being processed in the medical space.[63] Data access should also be restricted on a need to know basis.[64] The adoption of privacy by design as well as data protection impact assessments[65] are

56 In addition to the issues that have been discussed under this head, the European Data Protection Board (EDPB) has addressed the relationship between the GDPR and the Clinical Trial regulation. See, EDPB, The Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR) <https://edpb.europa.eu/our-work-tools/our-documents/avis-art-70/opinion-32019-concerning-questions-and-answers-interplay_es> accessed 29 October 2019.

57 Art 22 GDPR.

58 (n 47).

59 Art 22 (3) GDPR.

60 European Commission, 'Building Trust in Human-Centric Artificial Intelligence' COM(2019) 168 final.

61 For further readings on the need for human oversight and intervention for automated decision making, see, Peter Carey, *Data Protection A Practical Guide to UK and EU Law* (OUP 2018) 149 - 151.

62 Art 32(1) (a)-(d) GDPR. For further readings on the security of personal data under the GDPR, see, Stewart Room, 'Security of Personal Data', in Eduardo Ustaran (eds), *European Data Protection Law and Practice* (IAPP 2018) 169 -193.

63 Art 42 and 43 GDPR.

64 Art 5 (1) (f) FDPR.

65 For further reading, see: David Flaherty, 'Privacy Impact Assessments : An Essential Tool for Data Protection' (2000) 7 Privacy Law and Policy Reporter 5, 85.

practices which could also help ensure that appropriate security measures as well as relevant restriction of data access are adopted in processing activities of this nature.

4. Discrimination of Data Subjects

In order to tackle the (unlawful) discrimination that data subjects could potentially suffer in the hands of AI, it is important to ensure that the machine learning stage is inclusive of divergent data categories that covers people and cultures from different parts of the world. Some form of standardization will also be necessary so as to ensure that AI developers are bound by certain codes of conduct which include anti-discrimination practices in the machine learning stage. Most importantly, ethical considerations which will ensure that AI is all inclusive for people from all walks of life irrespective of ethnic, cultural, economic or other differences must be enshrined in relevant AI systems. In order to eliminate the (unlawful) discrimination of data subjects by AI systems, it is also recommended that AI systems should be made accessible for public review and auditing. This will allow interest groups and researchers to audit AI systems for discriminatory practices which could help limit the continued use of discriminatory AI.[66] It was an access to audit an algorithm of this nature that gave Obermeyer et. al. access to the widely used discriminatory algorithm in the USA that was discriminating against people of colour.[67] However, in granting such access to interest groups and researchers, it must be ensured that there is no access to the personal data of data subjects. This could be achieved by ensuring that fictitious data of a type that would usually be inputted into or used by the AI system is used for the audit purpose. This would invariably prevent the attendant risks that could arise from the use of actual (sensitive) personal data.

5. Data Transfers

In the event of data transfers outside the EU/EEA (particularly during telerobotic surgeries), the rules that apply under the extant provisions of the GDPR must be complied with. In such a case, the adequacy status of such a country will be one of the foremost considerations therein.[68] It is proposed that in accordance with the chorus of standardization that has been reechoed throughout this article, the adoption of AI-industry wide codes of conduct may be a very probable way of lawfully transferring personal data in the use of AI in medical practice.[69] Such code of conduct will help regulate the entire life cycle of the processing operation as anticipated under the provisions of the GDPR.[70]

From the specific recommendations that have been made above, one common denominator therein is the need to have more organization and structure in the development of AI particularly as same pertains to the practice of medicine. This is with the intention of creating a better system for the accountability of AI developers. In making a case for such structure in the use and development of AI, Mittelstadt[71] makes a distinction between the relationship between AI developers and data subjects in the use and regulation of AI on one hand and the medical relationship between a medical doctor and his patient on the other hand. Though there may be discrepancies on certain points, the practice of medicine is highly regulated with doctors bound to common aims, values, and trainings which is focused on promoting the interests of the patient, an attribute that is lacking in the development of AI. It is therefore necessary to ensure that AI systems that will be used in the practice of medicine are able to meet those common aims, values, and trainings that have been

66 For further reading on algorithmic auditing, see: Christian Sandvig et al 'Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms' (2014) <http://www-personal.umich.edu/~csandvig/research/Auditing%20Algorithms%20--%20Sandvig%20--%20ICA%202014%20Data%20and%20Discrimination%20Preconference.pdf> accessed 15 October 2019.

67 (n 51).

68 The adequacy decision is a decision of the European Commission certifying that a particular country meets the requirements for the processing of the personal data of EU residents. See, 'Adequacy Decisions: How the EU Determines if a non-EU Country has an Adequate Level of Data Protection' <https://ec.europa.eu/info/law/law-topic/data-protection/international-dimension-data-protection/adequacy-decisions_en> accessed 20 October 2019. For further conditions for transferring personal data outside the EU/EEA, please see Art 44-49 GDPR.

69 Art 40 GDPR.

70 For further readings on data transfers in the EU/EEA, see: Gregory Voss, 'European Union Data Privacy Law Reform: General Data Protection Regulation, Privacy Shield, and the Right to Delisting' (2017) 72 Business Lawyer 1, 221-233.

71 Brent Mittelstadt, 'Principles Alone Cannot Guarantee Ethical AI' (Nature Machine Intelligence, November 2019). <https://ssrn.com/abstract=3391293> accessed 28/10/2019

characteristic of the medical profession for ages. An example of such aims, values and trainings applicable to and binding on medical professionals are premised on the Hippocratic oath to which medical doctors pledge their allegiance.[72] In order to achieve a similar standard in the usage and development of AI in medicine, it is recommended that a regulatory body that will oversee the usage of AI in medical practice should be established such regulatory body will ensure compliance with the relevant data protection principles so as to ensure that compliance with relevant data protection laws (as well as other relevant fields of law) is a priority in the development and use of AI in medicine. Such regulatory body could even go as far as licensing developers of professional AI so as to ensure that certain values and principles are enshrined in AI development and to make the oversight and sanction for poor AI development a clearly established and defined process.[73] To this end, a Hippocratic oath which could regulate the development of technology and which is premised on 'proactively understanding the ethical implications of technology for all stakeholders, telling the truth about the capabilities, advantages and disadvantages of a technology, and acting responsibly in situations which are morally challenging' have been recommended and may be a good starting point for developing a code of ethics of some sort for medical technology.[74]

Another important consideration in the protection of (sensitive) personal data being processed by AI in the medical space is the balancing and consideration of all laws that regulate the practice of medicine. Some of such legislations include the earlier referenced 'medical devices directive',[75] and the 'directive setting standards of quality and safety for the processing of human blood and blood components'.[76] For instance, Article 24 of the latter directive[77] provides for data protection clauses which to a certain degree complements some of the data protection clauses encapsulated in the GDPR.

V. Conclusion

From the discussions in this article, AI development particularly as it pertains to medicine cannot be left in the hands of inventors and multinational companies alone. If the right of data subjects to data protection is to be engrained as a norm in the use of AI, it is important for regulators to start working towards the specific regulation of AI in the medical space particularly because of the relevance of this specific use of AI to the preservation of human life. While this article does not call for specific legislations to regulate the use of AI in the medical space, opinion and position papers by data protection supervisory authorities may go a long way in creating some certainty and soft-law specific to the use of AI in the medical space. As the usage of AI in medicine continues to develop, data controllers must also build themselves an ethical and moral compass so as to ensure that at no point is a derogation from the rights of people to data protection (and invariably, respect and dignity) an acceptable line to cross. This ethical and moral compass is however no substitute for the strict regulation of the use of AI in the medical space through some of the measures recommended above.

As the practice of medicine is a strictly regulated field, it is only logical that AI systems that will be used in the medical space will also be subject to equally effective forms of regulation that medical doctors are subjected to. This will in turn ensure that the use of AI in medicine does not have the counter effect of lowering the general standards applicable therein. In conclusion, despite the benefits being derived from AI and the glowing accolades it keeps receiving particularly in comparison to human beings (as is the case with medical doctors), AI must not be seen as a

72 Mark Rothstein, 'The Hippocratic Bargain and Health Information Technology' (2010) 38 Journal of Law, Medicine and Ethics 1.

73 Mittelstadt (n 71) 9.

74 Ali Abbas et al, 'A Hippocratic Oath for Technologists' (2018) Stanford University Graduate School of Business Research Paper No 19 - 4.

75 (n 29)

76 Directive 2002/98/EC of the European parliament and of the council of 27 January 2003 setting standards of quality and safety for the collection, testing, processing, storage and distribution of human blood and blood components and amending Directive 2001/83/EC.

77 The relevant provision provides among other things that 'Member States shall take all necessary measures to ensure that all data, including genetic information, collated within the scope of this Directive to which third parties have access have been rendered anonymous so that the donor is no longer identifiable. For that purpose, they shall ensure: (a) that data security measures are in place as well as safeguards against unauthorised data additions, deletions or modifications to donor files or deferral records, and transfer of information; (b) that procedures are in place to resolve data discrepancies; (c) that no unauthorised disclosure of such information occurs, whilst guaranteeing the traceability of donations.'

substitute to medical doctors but rather as an agent which makes the doctors' work easier. AI systems in their use in the medical space must always be made subject to the review of medical doctors to prevent some of the negative consequences of AI usage which have been discussed earlier in this article.