(A) Code of Conduct for Healthcare Data Practitioners

(B) The Research underlying the Design/formulation of the Code of conduct

2019-2020 Batch (Spring Semester)

Course: MSc in Business Analytics

Module: DATA GOVERNANCE AND ETHICS (IN6062)

Module Leader: Dr. Martin Cunneen

Submitted by:
Aishwarya C Inamdar
19084404

(A) Code of Conduct
for
Healthcare Data Practitioners

Table of Contents

Purpose

This Code of Conduct is based on the principles of securing the health data of the data subjects (i.e. patients) by fostering best practices using the proposed '**PLASTTICS'** framework (explained further). It states the statutory principles which the healthcare data practitioners should follow in order to ensure the safety, security, privacy and autonomy of the data of the public, whilst in their care, keeping their data the primary focus.

Scope

The Code applies to all the healthcare data practitioners in the Healthcare industry in order to ensure that the decisions made by the analysis of the health data are fair, transparent and consistent with supporting evidences [18].

How will this Code help the Healthcare Data Practitioners?

- This will offer a set of guidelines, so that the data practitioners can be assured of the standards they are expected to meet.
- They must utilise this Code to ensure that they work to the mark, and if not to adjust the way they operate and to identify areas for self-progress and enhancement in order to continue their professional development.
- Following this code ensures that the insights drawn from the data are represented without prejudice, the work is consistent and that the blemishes and biases are clearly revealed.
- This approach will help in the improvement of the health outcomes of the patients by tracking the current changing trends and predicting the future ones accordingly, thus, elevating patient engagement with predictive modelling and analysis based on the health data.

How will this Code help the public?

- The code is constructed on the fundamental concept of protecting the health data of the data subjects, and imparts an affirmation skeleton so that the patients can understand what ethical measures they can expect from the data practitioners.

Ethical Principles

This code encompasses the basic principles of ethical behaviour. The following basic principles of ethical behaviour should be followed by all the data practitioners:

- **P** - Privacy and Confidentiality
- **L** - Lawfulness
- **A** - Accountability
- **S** - Security
- **T** - Trust
- **T** - Transparency
- **I** - Integrity
- **C** - Communication
- **S** - Safety and Reliability

These principles are organised in the diagram below as the "**PLASTTICS**" framework:

Privacy and Confidentiality

Health data is highly sensitive. The data subjects reflect their trust by offering such extremely confidential information. According to the Data Protection Acts of 1988 and 2003 [3;4], the data subjects have a right to control when, where and with whom, to share their personal health information. A data practitioner should promise that any confidential health data collected unintentionally should be informed to the data subject. The third-party vendors or external stakeholders should also be guided about the various laws, practices and the GDPR guidelines regarding the safety of the patient's data thus assuring the privacy and confidentiality of the same. The data privacy laws are constantly altered and the retribution for non-compliance is acute [1]. The data practitioner should be constantly updated with the changing laws and policies.

Lawfulness

Lawful basis for processing personal data of user is necessary to protect vital information of the user, interests of patients and doctors and provision of health care. But, since the data is restricted to the user, data practitioners have to explicitly take a consent of the user to use the data or to forward it to "*Insurance Companies or Solicitors or Banks, and for other purposes which might not be obvious to the patient"* [13]. The practitioner must be able to demonstrate freely, the consent taken for the data in a clear and transparent manner. There are few divisions where the personal data of the user is necessary to be considered namely [14]:

- Medical History and Diagnosis.
- Providing health and social care
- Treatment analysis

- Managing health and social care systems for better services to patients.
- In-depth diagnosis for future use

The practitioner must also monitor all requests for elimination or withdrawals of consent, such as requests within the patient record and make sure that all removals are completed without undue delay [13].

Accountability

A data practitioner should be truthful with self and the data subjects. He/she should be accountable to provide a valid justification for his/ her decisions and actions and the repercussions of the same. He/she must function within the extent of their dominion always [5]. He/ she should be able to meet the legal commitments under the Data Protection and Freedom of Information Acts. They must also administer control over matters they are responsible for [5].

Security

- The data practitioner should apply encryption or pseudonymisation methodologies in order to prevent the exposure of sensitive health information.
- Audits on a routine basis should be performed for data security [13].
- Always keep the copies of data encrypted, so that in the cases of emergency, the server data can be deleted but the same copy will remain encrypted.
- Access controls is the baseline for every security aspect. Thus, a data practitioner should assign access controls like one user one id.
- Moreover, biometric scanners and access cards and keys should be used wherever possible.
- Data discovery and classification play an important supporting role in security by ensuring the sensitive data to be identified and tagged to receive proper level of protection.

HIPAA and GDPR follows the rule as stated, *"Focuses on securing the creation, use, receipt, and maintenance of electronic personal health information by HIPAA-covered organizations. The Security Rule sets guidelines and standards for administrative, physical, and technical handling of personal health information."* [15]

Trust

Handling the data and dealing with it in a proportionate form with the user's perspective is considered to be one of the main principles in the code of conduct. Since collecting the data can be crucial, handling and using the available data for public and individual benefit needs a supporting evidence. There are few norms by which a data practitioner can gain the trust [7;8]:

- Using the data anonymously: The data practitioners should make sure that they are making the data de-identifiable to the greatest degree as possible. If the data is released publicly, as it happens most of the time in medicine, it should be anonymised to the greatest degree so that the it will be out of data protection laws which will not be considered as personal data.
- Speaking of advance technologies, now-a-days the sensitive data is kept hidden under the protection act but in order to use the data, data practitioners can replace the real sensitive data with the plausible values termed as synthetic data.

- Seek permission: Under the acts for preserving the data, user or patient has right to choose whether the data they are giving can be used for further practices or not. Taking the example of Alexa, a popular AI of Amazon, which had medical records of the people although it had not been HIPAA certified. HIPAA is an Information Governance framework for medical data in the U.S. Since then, data practitioners are more aware for the user's privacy and trust for the data they keep.
- Trust and Transparency are the same sides of a coin, i.e. if the data of an individual or a group is obtained from the third-party resources, you should be clear and transparent about the data. As listed in article 13 and 14 of GDPR, the controller or the practitioner should enlist why, what and how the data is used in their context [6].

Transparency

The principle of transparency according to recital 58 of GDPR [9] states that, *"The principle of transparency requires that any information addressed to the public or to the data subject be concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualisation be used."*

- Due to the technological complexity, the data is available in many anonymous and third-party forms, makes the user incapable of understanding the data provided. The principle abides the data practitioner to define the whom and what purpose of the data, related to the individual, is collected.
- If the data collected by the practitioner is not clear and concise, they can request for more information regarding the individual/ subject in order to confirm the identity of the data subject.

According to the rights prescribed by GDPR [10], *"The controller should be obliged to respond to requests from the data subject without undue delay and at the latest within one month and to give reasons where the controller does not intend to comply with any such requests."*

- Moreover, the electronical transferred data should be presented with machine readable icons so that with the use of those machines, the data subject can access the data without haste.

Integrity

The accuracy, internal quality, and reliability of data are referred as data integrity.

- In order to promote high standards of patient care, data integrity preservation by the data practitioner is mandatory. In order to ensure the maintenance of data integrity, the data practitioner should assess the quality and uniqueness of the data as *'dirty data'* may include insufficient, absent or erroneous information which will lead to incorrect decision analysis thus hampering the quality of individual's healthcare [11].
- The source of the collection and creation of the health data should be traced so that it can be interpreted in an accurate manner and unhindered true insights can be gathered from it [1].

Communication [12]

- The data subjects must be informed of the intentions behind the collection of health data, its use and the amount of private information being made available to the external stakeholders/ clients.

- The data practitioner should ensure that the data subject has a clear visibility of their shared data, and also as per the laws and policies mentioned in GDPR and HIPAA, have the right to curb the flow of the personal information.
- The data subjects should be intimated of the financial transactions gained from the utility of their information along with the magnitude of such transactions.

Safety and Reliability

- The data produced by the technologies can be technically correct but since the data undergoes some augmentations, it should be clinically correct, which should be cross-checked with doctors. Since the data of images is large enough to separately label each image, many of AI technologies have helped the doctors in segregating those images. Data practitioner should help people understand the relevant aspects of a dataset's characteristics and origins which can help them better to understand the behaviour of models and systems involving that dataset [17].
- Clinical information assurance / care records assurance is accountable for keeping all the records safe and secure with highest quality of data possible.
- The information kept should be safe in terms of privacy and reliable in terms of accessibility as the quality of information impacts patient's safety as the information provided is based on clinical research [16].
- A data practitioner should always be aware of the risks to upgrade the data via unverified resources as well as should avoid multiple occurrences of same data over same server as it may result in checkpointing and loss of valuable data.
- He/she should always discuss issues of disclosure with a higher authority.
- Last but not the least, a data practitioner should always communicate through syntaxes wherever possible for internal communication, thus even though technology can be glitched, it will keep the data secured.

Code of Conduct Glossary [11; 18; 19; 20]

1) *"Data means a tangible or electronic record of raw (factual or non-factual) information (as measurements, statistics or information in numerical form that can be digitally transmitted or processed) used as a basis for reasoning, discussion, or calculation and must be processed or analysed to be meaningful."*

2) *"Data Practitioner means a professional who uses scientific methods to liberate and create meaning from raw data."*

3) *"Big Data means large data sets that have different properties from small data sets and requires special data science methods to differentiate signal from noise to extract meaning and requires special compute systems and power."*

4) *"Accountability is to be responsible for the decisions you make and answerable for your actions."*

5) *"Effective is to be successful in producing a desired or intended result."*

6) "*For "consent" to be valid, it must be given voluntarily by an appropriately informed person who has the capacity to consent to the intervention in question. This will be the patient, the person who uses health and care services or someone with parental responsibility for a person under the age of 18, someone authorised to do so under a Lasting Power of Attorney (LPA) or someone who has the authority to make treatment decisions as a court appointed deputy). Agreement where the person does not know what the intervention entails is not 'consent'."*

7) *"Transparent is to be open to public scrutiny."*

8) "*'pseudonymisation' means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person*."

9) "*'processing' means any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction."*

10) *"Data Integrity: The accuracy, internal quality, and reliability of data are frequently referred as data integrity."*

Bibliography

[1] Blast Analytics & Marketing. 2020. *Code Of Ethics - Our Promise As Analysts | Blast Analytics & Marketing*. [online] Available at: https://www.blastanalytics.com/code-of-ethics [Accessed 28 May 2020].

[2] Healthgrades | Evariant. 2020. *What Is Patient Engagement? | Evariant: The Leading Healthcare CRM Solution*. [online] Available at: https://www.evariant.com/faq/why-is-healthcare-data-management-important [Accessed 30 May 2020].

[3] *The Data Protection Act* 1988.

[4] *The Data Protection (Amendment) Act* 2003.

[5] WHO (2017) 'Code of Ethics and Professional Conduct - April 2017', (April). Available at: https://www.who.int/about/ethics/code_of_ethics_full_version.pdf.

[6] Gdpr-info.eu. 2020. [online] Available at: https://gdpr-info.eu/art-13-gdpr/ [Accessed 29 May 2020].

[7] Medium. 2020. *Is It Time For A Data Scientist Code Of Ethics?*. [online] Available at: https://towardsdatascience.com/is-it-time-for-a-data-scientist-code-of-ethics-210b4f987a8 [Accessed 29 May 2020].

[8] Data Science Association (2019) 'Data Science Code of Professional Conduct', *Dsa*, (i), pp. 1–12. Available at: https://www.datascienceassn.org/code-of-conduct.html.

[9] Gdpr-info.eu. 2020. [online] Available at: https://gdpr-info.eu/recitals/no-58/ [Accessed 29 May 2020].

[10] Gdpr-info.eu. 2020. [online] Available at: https://gdpr-info.eu/recitals/no-59/ [Accessed 29 May 2020].

[11] Vimalachandran, P. *et al.* (2018) 'Ensuring data integrity in electronic health records: A quality health care implication', *2016 International Conference on Orange Technologies, ICOT 2016*, 2018-January, pp. 20–27. doi: 10.1109/ICOT.2016.8278970.

[12] AI & Big Data Expo - Conference and Exhibition. 2020. *Ethics And Responsibility Within The Data Age | AI & Big Data Expo*. [online] Available at: https://www.ai-expo.net/thedataage/ [Accessed 29 May 2020].

[13] 'Processing of Patient Personal Data : A Guideline for General Authors : ICGP Data Protection Working Group' (2018), (October).

[14] Hse.ie. 2020. [online] Available at: https://www.hse.ie/eng/gdpr/gdpr-faq/hse-gdpr-faqs-public.pdf [Accessed 28 May 2020].

[15] Lord, N., 2018. *Healthcare Cybersecurity: Tips For Securing Private Health Data*. [online] Digital Guardian. Available at: https://digitalguardian.com/blog/healthcare-cybersecurity-tips-securing-private-health-data [Accessed 28 May 2020].

[16] HIQA (2010) 'An "As Is" Analysis of Information Governance in Health and Social Care Settings in Ireland', (January), pp. 1–73.

[17] Microsoft (2018) ‘How responsible innovation can lead to a healthier society’, (December). Available at: https://www.digitaleurope.org/wp/wp-content/uploads/2019/02/Healthcare-AI-Data-Ethics-2030-vision.pdf.

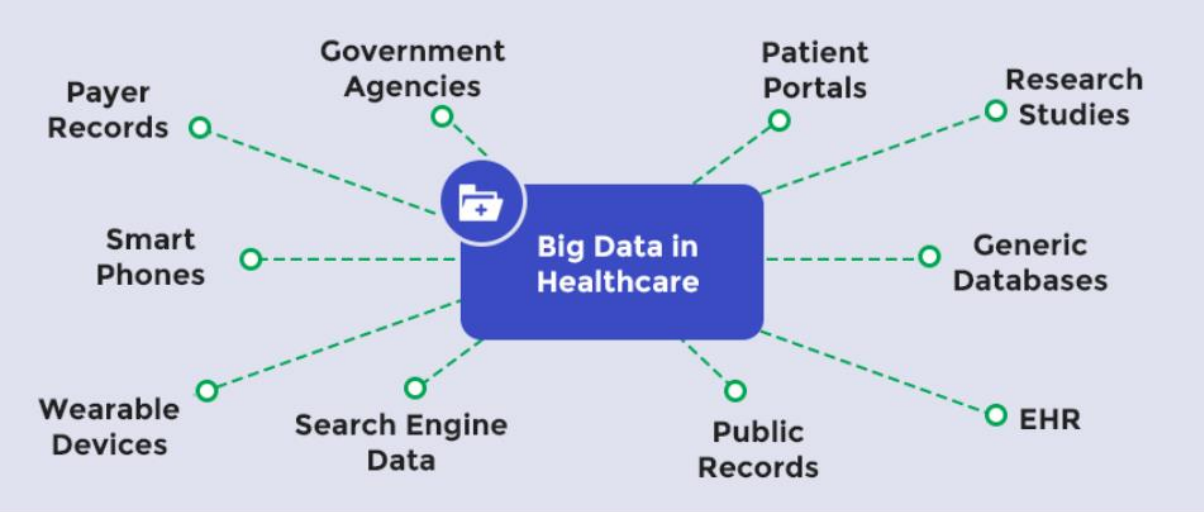
[18] Department of Health (2013) ‘Code of Conduct for Healthcare Support Workers and Adult Social Care Workers in England’, *Skills for Care*, p. 13. Available at: https://www.skillsforhealth.org.uk/images/services/code-of-conduct/Code of Conduct Healthcare Support.pdf.

[19] Data Science Association (2019) ‘Data Science Code of Professional Conduct’, *Dsa*, (i), pp. 1–12. Available at: https://www.datascienceassn.org/code-of-conduct.html.

[20] Group, W. (2013) ‘Processing of personal data’, 1(October), p. 2013.

(B) The Research underlying the Design/formulation of the Code of conduct

Technology is never unbiased because of its robustness. Being a robust orientation, it has an intrinsic tendency to shape human lives and improve them. The new forms such as high-speed internet, Big data, Cloud computing, Internet of Things, smart applications and appliances, Artificial Intelligence are all the biggest human achievements and developments which enables the capacity of a machine to process more than needed, achieving more than required and satisfy the impossible. Profoundly, these technological advancements are more to the improvement side in health care giving rise to the concept of '*Intelligent Health*' thus giving the healthcare providers (Doctors) an in-depth information of the patients (Sinhasane, 2019).



How to Apply Big Data in Healthcare – source: https://mobisoftinfotech.com/resources/blog/big-data-in-healthcare-industry/

Networked data is increasing in demand but also posing a threat to the adoption of innovative healthcare technologies by the introduction of '*dirty data*' (Vimalachandran.P, 2018). Technological feasibility on one side is proving to excel in handling the patients for all kinds of diseases but on the other side ethical expectations comes with a cost in privacy and data governance. Even with advances in data protection, health data lacks availability and accessibility for researchers, patients and doctors when they need it to help realize better outcomes (Microsoft, 2018).

The General Data Protection Act (2018) and Medical Practice states that a constructive administration using the six privacy principles of *lawfulness, fairness & transparency, purpose limitation, data minimization, accuracy, storage limitation and integrity & confidentiality* should be imposed for ensuring the control of the data subjects over their personal information and giving them the authority to *access and be acknowledged about the processing of their data, make necessary changes in case of inaccuracy, 'right to be forgotten', curb the information processing, permission for data portability and a choice of not being subjected to automated decision making and profiling* (O'Shea.D, 2019). According to the National Health Information Strategy (NHIS) of IRELAND, 2004, there is a framework which is responsible for organizations and individuals which ensures them that their personal information and private health data is handled securely, carefully and with confidence to deliver the best possible care. The framework is Information Governance which states the definition according to NHIS (HIQA, 2010) that, *"A strategic framework that brings coherence and transparency to information initiatives and which is responsive to the spectrum of issues and concerns of*

those involved. Issues such as information sharing, health surveillance, quality assurance, confidentiality, privacy records management, freedom of information and data protection are included". The theme 8 of *National Standards for Safer Better Healthcare* published by the Health Information and Quality Authority (HIQA) states that *good information governance enables services and individuals to ensure all information, including personal information, is handled securely, efficiently, effectively and in line with legislation* (HIQA, 2012). The *Guidance on information governance for health and social care services in Ireland* aims at supporting the data practitioners with the collection, segregation, analysis and sharing of personal information in a lawful and in a productive and efficient manner (HIQA 2010, 2012).

Referring the code of conducts of various healthcare facilities and organizations like HSE, Microsoft in Healthcare, Mount Sinai Hospital, WHO, etc. and following the data compliance policies, frameworks (GDPR, HIPAA, IG)and regulatory acts (The Data Protection Act) proposed by governments of various regions, I have and proposed the code of conduct for the healthcare data practitioners using the **'PLASTTICS'** framework that is constructed by identifying the key principles stated in the aforementioned documents. The framework is proposed by maintaining the privacy of the information of the data subjects as the key focus, with a view of delivering high quality healthcare & support and an unbiased delivery of decisions on processing and analyzing high quality data. The underlying research of the chosen principles is as mentioned below:

- **Privacy and confidentiality (Healthcare Solutions | Zettaset, n.d.):** As the data generated for the healthcare is huge, big data is an essential tool to manage the data with all the privacy principles for a data practitioner. With the inaugurations of HIPAA, GDPR and HITECH, the data is electronically surfed over the internet. Thus, maintaining the privacy is essential as healthcare data is most sensitive for a patient. With the increase in healthcare data electronically, there comes a risk factor of breaching the data by cyber hackers as well as misusing the data. The common example as stated in the code of conduct is the Alexa's encryption under the norms of HIPAA infrastructure. Keeping the privacy issues minimal and gaining the trust of patients or data subjects, data practitioners are abided by the principal or privacy policy for each institution or organization.
- **Lawfulness:** Healthcare data being the most sensitive kind of data must be processed or shared with external stakeholders/ clients under lawful grounds in order to avoid data breaches.
- **Accountability (ICGP Data, 2018):** Every data has an accountability principle which has to be followed under the protections act given by GDPR. According to the act, the controller or data practitioner has a requirement to keep certain records and those include: "*Regular Information Security Audits"* (RISA), confidentiality agreements, "*Records for Processing activities"*, consensual records, records for training and awareness and etc.
- **Safety and Reliability (Nass, Levit, Gostin and Rule, 2020):** Personal data can be extremely sensitive and potentially embarrassing at times. Any kind of health research or analysis requires data in huge quantities. In case of security breach in such a case, there are chances of potential harm being caused to the data subject. For eg, the disclosure of a data subject affected with some sexually transmitted disease can cause

psychological and public damage. Thus safety and reliability of an individual's data is a must.

- **Trust (Microsoft, 2018):** Considering all the risks and vulnerabilities associated with the patient's data, considering communication as a key in patient's trust, data practitioners are responsible to maintain the data with all the anonymous practices. Keeping the data de-identified, not only helps the doctors but also the researchers to maintain the provisions under GDPR amendment laws to ensure innovative research projects are preserved. New innovations in the field of medical and healthcare technology promotes new models of research to make the data more easily available for the user, thus gaining more trust from the user. Securing the data with new solutions proposed can enable secure machine learning with multiple data/ sources, hence reducing the chance of duplicity from the same source.
- **Transparency (Transparency, n.d.):** The term *transparency* is a principle which allows the user to trust the institution and data practitioners. This principle abides the controller to demonstrate the data with utmost opaqueness and follow the guidelines with all the integrity. As stated in GDPR, transparency is a requirement irrespective of the legal bias throughout the lifecycle of medicine. The requirements for transparency provided to individuals must comply with:

❖ Concise and intelligible content.
❖ Easily accessible data.
❖ The requirement for clear and plain language is of particular importance when providing information to children.
❖ Must be free of charge.

The articles stated in GDPR articles 12,13,14 and recites 39 and 58 give a simple understanding about the data transparency compliance and the PLASTTICS framework have a brief introduction and conduct for data practitioners for transparency.

- **Integrity (Vimalachandran.P, 2018): "***Integrity refers to consistent, accurate and significant data"* which is a big concern in healthcare industry. Failing to maintain this could lead to incompleteness, medical errors and poor quality of data. Thus, data integrity is a common practice by data practitioners to keep the data maintained electronically. A common example in this is naming conventions for the files of same patient and inconsistences between the data fields. Thus, an extra value identifier or structuring the data can help the data practitioners to maintain the integrity and can cause minimum errors.
- **Communication:** Communication in healthcare is referred to as the sharing of the information to the clients or the external stakeholders with the consent of the data subjects. Under the GDPR, the consumers have the right to be informed about the details where their data is being shared and for what purpose.
- **Security:** Manual data as well as electronic data has to be secured from the physical damages as well as cyber damages. The most common form of electronic security is encryption and decryption. Since, the data travels through internet, the insecure server is a major threat and thus frequent audits have to be done in order to maintain the integrity of data.

  There are main 3 levels of security in health care industry which is included in the guideline of code of conduct namely (What are the 3 Key Layers in Healthcare Data Security?, n.d.):

- Physical security: Data Centers and physical cloud
- Logical Security: Virtualized system, anti-virus, malware detector, VPN, Identity awareness etc.
- Compliance: PCI, HIPAA, AWS, Safe Harbor and etc.

Following the PLASTTICS framework in the healthcare organization will definitely ensure secured processing and analysis of the sensitive health information of the individuals making the data robust and thus help in making informed decision.

References

1) ‘Processing of Patient Personal Data : A Guideline for General Authors : ICGP Data Protection Working Group’ (2018), (October).

2) HealthITSecurity. n.d. *What Are The 3 Key Layers In Healthcare Data Security?*. [online] Available at: https://healthitsecurity.com/news/what-are-the-3-key-layers-in-healthcare-data-securit [Accessed 29 May 2020].

3) HIQA (2010) ‘An “As Is” Analysis of Information Governance in Health and Social Care Settings in Ireland’, (January), pp. 1–73.

4) HIQA (2012) ‘Guidance on information governance for health and social care services in Ireland’. Available at: https://www.hiqa.ie/system/files/Guidance-on-information-governance.pdf.

5) Hrb.ie. n.d. *Transparency*. [online] Available at: https://www.hrb.ie/funding/gdpr-guidance-for-researchers/gdpr-overview/gdpr-principles/transparency/ [Accessed 29 May 2020].

6) Microsoft (2018) ‘How responsible innovation can lead to a healthier society’, (December). Available at: https://www.digitaleurope.org/wp/wp-content/uploads/2019/02/Healthcare-AI-Data-Ethics-2030-vision.pdf.

7) Nass, S., Levit, L., Gostin, L. and Rule, I., 2020. *The Value And Importance Of Health Information Privacy*. [online] Ncbi.nlm.nih.gov. Available at: https://www.ncbi.nlm.nih.gov/books/NBK9579/#:~:text=SECURITY%20OF%20HEALTH%20DATA,be%20sensitive%20and%20potentially%20embarrassing. [Accessed 28 May 2020].

8) O’Shea, D. (2019) ‘General Data Protection Regulation and Medical Practice’. Available at: https://rcpi-live-cdn.s3.amazonaws.com/wp-content/uploads/2019/03/GDPR-Guidance-Document-March-2019.pdf.

9) Sinhasane, S., 2019. *How Big Data Is Changing The Healthcare Industry?*. [online] Mobisoftinfotech.com. Available at: https://mobisoftinfotech.com/resources/blog/big-data-in-healthcare-industry/ [Accessed 29 May 2020].

10) *The Data Protection (Amendment) Act* 2003.

11) *The Data Protection Act* 1988.

12) Vimalachandran, P. *et al.* (2018) ‘Ensuring data integrity in electronic health records: A quality health care implication’, *2016 International Conference on Orange Technologies, ICOT 2016*, 2018-January, pp. 20–27. doi: 10.1109/ICOT.2016.8278970.

13) WHO (2017) ‘Code of Ethics and Professional Conduct - April 2017’, (April). Available at: https://www.who.int/about/ethics/code_of_ethics_full_version.pdf.

14) Zettaset. n.d. *Healthcare Solutions | Zettaset*. [online] Available at: https://www.zettaset.com/solutions/data-privacy-protection-healthcare/ [Accessed 28 May 2020].