HOW DOES DATA PROTECTION REGULATION ADDRESS CONTEXT TRANSGRESSION IN THE CASE OF "BIG DATA" PROCESSING OF ELECTRONIC HEALTH RECORDS, AND HOW EFFICIENTLY?

Master's thesis

Submitted in partial fulfillment of the requirements for the degree of

LL.M. Law and Technology

Supervisor: Dr. Linnet Taylor

Second reader: Maša Galič

Tilburg University

Tilburg Law School – Tilburg Institute for Law, Technology and Society (TILT)

Academic Year 2017-2018

Table of Contents

CHAPTER 1: INTRODUCTION

1.1. PROBLEM DESCRIPTION AND SIGNIFICANCE

Rapid technological developments, increased data storage capabilities and reduced costs have made tech giants looking to diversify their business activities turn their attention towards the application of advanced data analytics to electronic patient records for various purposes. Although there have already been a number of health record processing partnerships in the USA, eHealth has comparatively recently gained momentum in Europe. Assessment of the changes in the data protection framework of the European Union and the United Kingdom seems timely. This thesis will therefore examine the legal response through data protection to context transgression in the case of "big data" processing by private companies partnering with healthcare providers, its impact on the stakeholders in the health domain and thus its efficiency.

There is no single, universally accepted definition of "Big Data", as the scope and meaning attached to the phrase evolves along with technological advances and conflicting interpretations[1]. However, one of the most popular definitions is provided by the Gartner IT glossary, and describes big data as "…high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making." [2] Most of the key terms in business and innovation nowadays are all associated with such high-volume datasets and the advanced techniques for extracting knowledge from them: "Big Data" analytics, artificial intelligence (AI), machine learning, data mining, algorithmic profiling *etc.*

Data science has undoubtedly revolutionized commerce and marketing, so there is also a lot of optimism with regards to its potential for medical care, research and development.[3] There is considerable interest worldwide in implementing eHealth services (the application of Information and Communication Technologies for the provision, management and

[1] For a discussion of "big data"-related definitions and their scope, see *e.g.* Amir Gandomi, Murtaza Haider, 'Beyond the hype: Big data concepts, methods, and analytics', International Journal of Information Management, (2015), 35, 137–144.
[2] Gartner IT glossary 'Big Data' <http://www.gartner.com/it-glossary/big-data> Last accessed on 31 January 2018.
[3] See, *e.g.* Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century', (Communication) COM (2012) 0736 final <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012DC0736> Last accessed on 4 December 2017.

enhancement of healthcare systems[4]) in order to improve efficiency and reduce costs. While the number of projects utilizing data collected by tech giants through social media, fitness devices or apps for eHealth purposes is growing, there are justified concerns about the quality of such data[5] and thus the conclusions based on their analysis,[6] which some medical researchers consider entirely invalid.[7] On the other hand, data collected by healthcare providers are considered to be more accurate on account of being recorded under the supervision of medical staff, so algorithm-driven research would, in theory, have a better scientific basis if patient records were used. Health records of individuals, in digital form, have thus become central to most health informatics applications.[8]

It is therefore no surprise that more and more ICT companies are turning to Electronic Health Records (EHRs) as the most viable source of valuable insights. One of the latest examples of this is Apple, the world's largest (in terms of revenue) ICT company,[9] which recently "made a jump into the electronic health records game by allowing patients to aggregate their records on their iPhones",[10] and plans to develop further innovative uses of the EHR data shared voluntarily by iPhone owners.[11]

It should be noted that there are different kinds of Electronic Health Records. In Europe, for example, the Article 29 Working Party (the EU data protection expert advisory body which has now been replaced by the European Data Protection Board) has distinguished three types of

---

[4] Diane Whitehouse, Carlisle George, and Penny Duquenoy, 'eHealth: legal, ethical and governance challenges - an overview', Global Telemedicine and eHealth Updates: Knowledge Resources Vol. 4, (2011, International Society for Telemedicine & eHealth (ISfTeH)), 423-428.
[5] Tamar Sharon, 'The Googlization of health research: from disruptive innovation to disruptive ethics', (2016) Personalized Medicine, Volume 13, Issue 613 Oct 2016, 4.
[6] A prominent example of inaccurate conclusions resulting from analysis of incomplete data is Google's failed Flu Trends project. See Adam Kucharski, 'Google's flu fail shows the problem with big data', (The Conversation, 24 October 2013) <https://theconversation.com/googles-flu-fail-shows-the-problem-with-big-data-19363> Last accessed on 4 December 2017.
[7] Sally Wyatt, Anna Harris, Samantha Adams, Susan E Kelly, 'Illness Online: Self-reported Data and Questions of Trust in Medical and Social Research', (2013) Theory, Culture & Society, vol. 30, Issue 4, 131 – 150, 1.
[8] Pradeep K. Sinha et al., *Electronic Health Record: Standards, Coding Systems, Frameworks, and Infrastructures* (1st edn., Wiley 2013) 3.
[9] Fortune 500, <http://fortune.com/global500/apple/> Last accessed on 2 February 2018.
[10] Rachel Z. Arndt, 'Apple is officially in the EHR business. Now what?' (*Modern Healthcare*, 26 January 2018), <http://www.modernhealthcare.com/article/20180126/NEWS/180129910> Last accessed on 2 February 2018.
[11] *Ibid.*

EHRs: decentralised, centralised and patient-controlled.[12] The types of information included in the record can also vary, depending on the national healthcare system or the particular provider, and, correspondingly, so can the consequences of using and re-using the data therein. The defining characteristics of any EHR according to the Working Party are as follows:

"A comprehensive medical record or similar documentation of the past and present physical and mental state of health of an individual in electronic form and providing for ready availability of these data for medical treatment and other closely related purposes."[13]

It is easy to see from the description above why recent health data-sharing agreements between public bodies and private "Big Data" corporations have brought significant ethical and legal issues to the fore of public and academic opinion. The most prominent example is a collaboration which began in 2015 between Google's British-based artificial intelligence (AI) subsidiary, DeepMind Technologies Limited, and the Royal Free London NHS Foundation Trust. It involved the transfer of identifiable patient records from the public hospitals to the private corporations without explicit consent, for the purpose of developing a clinical alert app (called Streams) for acute kidney injury (AKI).[14] The British medical information governance rules require obtaining the explicit consent of each patient, and apply in all cases where identifiable health data is passed on to a third party which is not in a direct care relationship with that patient.[15] The main legal issue regarding Streams concerned the reliance on the "direct care" exemption by the contracting parties. They argued that the transfers and the processing of the highly sensitive health data of all of Royal Free's patients by DeepMind were carried out for the purposes of preventing AKI, therefore for direct care. As such, instead of explicit consent, they claimed that they had the "implied consent" of all patients, including those who were not monitored for AKI because they had not had the prerequisite renal blood test, and

[12] Article 29 Working Party, Working Paper nr 131, "Working Document on the processing of personal data relating to health in electronic health records (EHR)", adopted on 15 February 2007, 17: with regard to the third alternative, the Art. 29 Working Party refers to the French system.
[13] *Ibid*, 14.
[14] See Julia Powles and Hal Hodson, "Google DeepMind and healthcare in an age of algorithms" (2017) Health and Technology, December 2017, Volume 7, Issue 4, 351–367< https://doi.org/10.1007/s12553-017-0179-1> Last accessed on 4 December 2017.
[15] *Ibid*, at 2.2.

those who had ended their direct care at the London hospital.[16] After investigating the matter, the UK's Information Commissioner sent a letter to Royal Free stating her findings and, *inter alia*, advising that "implied consent" constituted an inappropriate legal justification and thus a breach of data protection law.[17] Data protection and individual privacy concerns were at the heart of the public debate.[18] Academics also questioned the lack of transparency about the precise terms of the deal and the risk of context transgression,[19] the lack of engagement with patients[20] or the risk of Google and DeepMind gaining "undue and anticompetitive leverage over the NHS".[21]

The collaboration between Google DeepMind and Royal Free is not the first or only one of its kind, but it demonstrates the general trend. According to news reports, the e-commerce giant Amazon aims to pursue similar patient-record sharing agreements with healthcare institutions as those concluded by Google.[22] After months of speculation on Amazon's ambitions in the healthcare domain, especially in respect of health records,[23] rumours about the application of advanced analytics and machine learning technologies to health records still persist.[24] Furthermore, Amazon's plans, reported last year,[25] to simultaneously launch a pharmaceutical

---

[16] *Ibid*.
[17] UK Information Commissioner's Office, Letter dated 3 July 2017 to Sir David Sloman, Chief Executive of Royal Free NHS Foundation Trust < https://ico.org.uk/media/action-weve-taken/undertakings/2014353/undertaking-cover-letter-revised-04072017-to-first-person.pdf> Last accessed on 4 August 2018.
[18] See, *e.g.* 'Google DeepMind NHS app test broke UK privacy law', *BBC news*, <http://www.bbc.co.uk/news/technology-40483202> Last accessed on 12 October 2017; 'Google's Deepmind NHS deal 'inexcusable', says academic paper', *The Register* (16 March 2017), < https://www.theregister.co.uk/2017/03/16/googles_deepmind_and_royal_free_hospital_deal_inexcusable/> Last accessed on 12 October 2017.
[19] Powels & Hodson (N14) at 2.1 and 4.4.
[20] *Ibid*, at 4.1.
[21] *Ibid*, at 3.2.
[22] *Ibid*.
[23] Sejuti Banerjea, 'Amazon on the Brink of EHR Deal with Cerner', *Nasdaq*, (29 November, 2017), <http://www.nasdaq.com/article/amazon-on-the-brink-of-ehr-deal-with-cerner-cm884382> Last accessed on 2 February 2018.
[24] Annie Palmer, 'Amazon's secret health lab revealed: 'Grand Challenge' working on everything from curing cancer to using AI to analyse medical records', *Daily Mail* (5 June 2018) <http://www.dailymail.co.uk/sciencetech/article-5810087/Amazons-secret-health-lab-revealed-Grand-Challenge-working-cancer-research-medical-records.html> Last accessed on 1 July 2018.
[25] See Stephanie Baum, 'Three perspectives on how Amazon could disrupt the pharmacy space', (*MedCity News*, 2 June 2017), <https://medcitynews.com/2017/06/three-perspectives-amazon-disrupt-pharmacy-space/?rf=1/> Last accessed on 4 December 2017.

offshoot in order to disrupt the pharmaceutical supply chain have recently come to fruition. The e-commerce giant acquired the small online pharmacy PillPack on 28 June 2018.[26] The argument has been made that this could have the effect of pushing the cost of medication down as a result of competition with traditional pharmacies.[27] There would undoubtedly be benefits in terms of the efficiency and speed of delivery of prescribed medication to patients, especially those with mobility difficulties. Google's or Amazon's experience with AI analytics and vast resources cannot be discounted either, neither can their potential for medical research and development.

Innovations notwithstanding, the multi-faceted business model of tech giants and their long-term private profit-making agendas pose problems when it comes to the application of "big data" processing of EHRs. As the lines between the status of a consumer and that of a patient (and/or clinical trial participant) become more and more blurred when data is digitally analysed for multiple purposes, there is inevitably a clash between the world of commerce and the domain of medicine. The threat of context transgression led to public distrust of the Google DeepMind/NHS collaboration. So why was it troublesome? According to the privacy theorist Helen Nissenbaum,[28] norms and privacy expectations differ depending on contextual circumstances: on the nature of the information, the type of relationship in which information is transferred and the uses to which it is put. The electronic health record, in essence, enables the processing of the information shared within the original context (traditional doctor-patient relationship governed by medical confidentiality and a specific set of ethical imperatives) by private companies (outside of that traditional relationship and not necessarily governed by the same rules). To complicate matters further, the methods utilised by these companies to discover patterns and correlations from datasets are often opaque to the general public and the data subjects especially, which leads to unrealistic privacy expectations or fears. Nevertheless, as any new technology, they can have unintended and unpredictable consequences. Lastly, the

[26] Bruce Japsen, 'It's Official: Amazon Enters Pharmacy Business With PillPack Acquisition' , (*Forbes*, 28 June 2018), < https://www.forbes.com/sites/brucejapsen/2018/06/28/its-official-amazon-enters-pharmacy-business-with-pillpack-deal/#6035c07113fa> Last accessed on 1 July 2018.
[27] *Ibid*.
[28] Helen Nissenbaum, *Privacy In Context:Technology, Policy, and the Integrity of Social Life*, (Stanford University Press, 2010).

profit-seeking nature of these corporations brings into question the actual and potential future uses of the EHRs and the knowledge derived from them. It is therefore prudent to examine the existing ethical and legal frameworks in Europe so as to determine the possible impact of context transgression on the key stakeholders in the healthcare sector.

1.2. LITERATURE REVIEW

There is already an impressive amount of literature on the legal and ethical concerns arising out of the collaborations between public health institutions and private tech corporations. It is unsurprising that it has grown exponentially in recent years with the increasing number of regional and state initiatives to implement eHealth services. Due to the vital importance of healthcare provision for the functioning of society, there is significant public interest in efficient and timely regulation of this field. Academics[29] have stressed the need for a comprehensive overhaul of the existing approaches in order to address ethical concerns regarding the rise of commercial AI services in the public health domain.

A major issue that has not yet received attention is how certain recent legal reforms in the EU may affect advanced analysis of EHRs by private tech corporations. In Europe, because of the ageing population and medical staff shortages, the national healthcare systems have an interest in using the help of commercial data science to improve efficiency. On the other hand, can and should the vast resources and analysis expertise of AI companies be utilized by public healthcare institutions, despite all the risks to the patients' rights and liberties? Some authors have already discussed certain gaps and inefficiencies in the old personal data protection regime, but the new legislation inevitably brings uncertainty and therefore requires close scrutiny.

While individual rights are the primary focus of academic ethics research, less consideration is given to group privacy and public policy implications. Some authors[30] discuss the threat of placing citizens or researchers at the mercy of tech corporation's profit-seeking agendas, but

[29] See *e.g.* Nuffield Council on Bioethics, *The Collection, Linking And Use Of Data In Biomedical Research And Health Care: Ethical Issues* (2015) <http://nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data_web.pdf> Last accessed on 16 February 2018; Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013) 6; Sharon (N5), 2.
[30] Powels & Hodson (N14); Sharon (N5) 2.

that is a bit of an oversimplification as it does not consider all other parties in the healthcare and research domain. Indeed, the many stakeholders involved – patients, tech companies, healthcare institutions, policy-makers; as well as the potentially affected by data mining: pharmaceutical companies, university researchers, insurance companies, *etc.* - all have a different legal and economic status. Regulation or lack thereof can place some of the actors in a strong position to benefit at the expense of others. It is, therefore, important to envisage all the stakeholders who would gain special advantages and all those who may be harmed,[31] in order to assess the potential new problems and strengthen the protection of certain rights. There will always be tension between the rights of all stakeholders, so concessions should be made on all sides, but to varying degrees.

It appears that not enough due consideration is given in literature to the balance between public and private rights and interests and the protection of legitimate pursuits and stakes of all concerned parties in the given context.[32] These should be identified for regulatory measures to be efficient. The recent European Commission consultation on eHealth was directed at "citizens, patient organisations, health and care professionals, public authorities, researchers, industries, investors and users of digital health tools".[33] The Commission needs to weigh the concerns of all interested parties, and so should academics. There is a legitimate public interest in technological development and research, particularly in the healthcare sector. Investment and innovation should not be stifled due to risk-aversion. It is indeed imperative to recognize the potential threats, but it is only the first step towards creating the pertinent safeguards.

1.3. THE RESEARCH QUESTION AND SUB-QUESTIONS

[31] This would be in line with the approach advocated by the author and cultural critic Neil Postman in his talk in Calvin collage, July 1998, "6 questions that needs to be addressed when anyone tells you about new technology".
[32] For example, there is a similar discussion of the stakeholders in the context of mHealth. See Petersen C, Adams SA, DeMuro PR., 'mHealth: Don't Forget All the Stakeholders in the Business Case.', (2015) Eysenbach G, ed. Medicine 20 2015; 4(2):e4. < https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4713907> Last accessed on 4 December 2017.
[33] European Commission,'Public Consultation on Health and Care in the Digital Single Market', (Strategy on Digital Market Policies – Consultation) <https://ec.europa.eu/digital-single-market/en/news/public-consultation-health-and-care-digital-single-market> Last accessed on 4 December 2017.

The thesis will examine how data protection law addresses the potential context transgression which may result from commercial "big data" processing of Electronic Health Records (EHRs) by tech companies.

In order to answer this central research question some relevant sub-questions will be addressed:

(1) What are the public and private interest conflicts stemming from the use of commercial data science by tech companies in the healthcare sector?
(2) What constitutes context transgression and what are the relevant ethical considerations regarding the risks of context transgression?
(3) What is the potential impact on the various stakeholders resulting from context transgression?
(4) How does data protection regulation in the UK, on the one hand, and the EU, on the other, address context transgression consequences and how efficiently? What are the roles of the principle of purpose limitation, patient autonomy, consent and control? Can pseudonymisation or anonymisation be efficient safeguards in the "big data" sphere?

1.4. METHODOLOGY

The methodology that the thesis will follow is that of traditional (doctrinal) legal research along with comparative legal research. Addressing each of the aforementioned sub-questions will contribute to answering the central research question and reaching conclusions regarding the regulatory gaps and challenges.

The starting point for the analysis will be the description of the intricate network of values and objectives of the different public or private entities and individuals in the healthcare domain. This will help delineate the potentially affected parties. Then the concept of context transgression will be examined in detail in order to illustrate the magnitude of the problems arising from it. Further, some examples of the adverse effects on each of the key actors and the ripple effect on the healthcare system and society will be discussed.

In order to assess the ways that stakeholders' interests can be protected from the negative aspects of context transgression, a comparative analysis of the regulatory safeguards and strategies of both the UK and the EU will follow. The contrast is on several levels: between national and supranational regulation, and between slightly diverging ethical perspectives.

There are valuable lessons to be learned from the UK's experience with the implementation of AI in the healthcare sector, as well as much debate surrounding it. On the other hand, there are legal challenges arising out of the multinational status of tech corporations that can only be addressed through supranational law. This would illustrate clearly whether regulation can be more efficient on the national or supranational level.

In essence, the main research question will be answered on the basis of a desk study analysis. The General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) which replaces the Data Protection Directive 95/46/EC (DPD), EU policy documents and consultations, Article 29 Working Part opinions will be evaluated in respect of the EU regulatory approach. To contrast them with the UK regulatory framework, the new UK Data Protection Act 2018, which replaced the Data Protection Act 1998, Information Commissioner's decisions, Nuffield Council on Bioethics[34] reports *etc.* will be discussed.

Finally, conclusions from the assessment of the regulatory responses to the risks and consequences of context transgression will be made to determine the efficiency of data protection measures in the UK and EU.

1.5. THESIS STRUCTURE

The thesis will proceed in three chapters and a conclusion. To begin with, the preliminary questions in Chapter 2 will serve as an introduction to the complicated landscape of the healthcare sector and its key actors. The tension between public and private interests will be explored to illustrate why context transgression is a very real possibility. Potential areas of conflict with the other concerned parties will be highlighted. The relevance of the context integrity theory to the EU and UK regulatory framework will be explained. Next, the potential impact on each of the main stakeholders will be discussed comprehensively in Chapter 3 and illustrated with examples. Further, in Chapter 4 the different regulatory approaches to data protection of electronic health records in the EU and the UK will be contrasted. The strengths and weaknesses of each will be elaborated on. The underpinning ethical principles and policies will also be emphasized. The extent to which such adverse consequences are addressed by the

[34] Nuffield Council on Bioethics (N29).

EU and the UK regulation will be analysed to demonstrate the efficiency of each regime, as well as to stress the gaps and weaknesses of the legislative instruments. The conclusion will summarise the main points of the comparative analysis of the limits of EU and UK data protection law as a regulatory tool to prevent context transgression. Lastly, an assessment will be made on that basis to highlight the strengths of each regulatory approach to balancing the rights of the stakeholders and their efficiency.

CHAPTER 2: PUBLIC AND PRIVATE INTERESTS. CONTEXT TRANSGRESSION

2.1. CONFLICTING INTERESTS

When it comes to partnerships between healthcare providers and tech companies involving innovative processing of healthcare data, there are many competing public and private interests. For example, there is a public interest in the responsible use of data to support scientific research, innovation and improvement of health services.[35] However, confidentiality of sensitive health data is traditionally deeply entrenched in codes of practice for medical staff, and is still expected by patients in most countries.[36] It remains a vital prerequisite for patient trust in the healthcare system or provider, as became apparent in the DeepMind/Royal Free media scandal mentioned above.

"Big data" processing methods have made possible the building of a strong evidence base for prediction, prevention and treatment of diseases, which is in the interests of public health. However, machine learning and data mining can challenge individual and group privacy by revealing unexpected links and patterns in patient records on which to base new medical hypotheses about the relevant data subjects. The utilisation of the new insights is the primary area of concern here: whether they are applied for the diagnosis and direct treatment of the individual patient, or to further medical research, or for some other secondary purpose. In the last scenario, in particular, the issues of patient autonomy, as well as free and informed consent, are especially pertinent. Depending on the specific secondary purpose of processing and whose vested interests it serves mainly, it may be very difficult to justify exposing patients to privacy-infringement related risks.

The push for personalised medicine, which in theory benefits patients, has also provided unprecedented commercial opportunities for the private IT companies with expertise in such processing and experience of monetizing the value of datasets. Although policy-makers are focused on increasing economic growth from the life sciences, and innovating companies have

[35] Nuffield Council on Bioethics (N29).
[36] See, *e.g.* Eleni Entzeridoua, Evgenia Markopouloua, Vasiliki Mollaki, 'Public and physician's expectations and ethical concerns about electronic health record: Benefits outweigh risks except for information security', (2018) International Journal of Medical Informatics, Volume 110, February 2018, 98.

a legitimate right to return of their research and development (R&D) investments, academic debate has highlighted the risk of "lock-in" situations, *i.e.* a private corporation gaining undue leverage over public health institutions using their products and services. In such situations, instead of reducing the cost of healthcare *per capita*, the opposite may happen. Nevertheless, healthcare providers and policy-makers are often persuaded by claims that AI as a tool to assist physicians can increase their efficiency and thus address medical staff deficits, and that secondary-purpose analysis can improve resource allocation, planning and quality control.[37]

On the basis of the expected potential of commercial advanced data analytics to revolutionize the healthcare sector, tech companies, such as IBM, have lobbied for self-regulation, arguing that a top-down regulatory approach would stifle innovation.[38] But the health domain is typically governed by comprehensive legislation, in addition to strict codes of practice, and to exclude certain commercial actors from their scope can adversely affect other stakeholders. Private or public entities, *e.g.* educational institutions, which conduct traditional medical research are restricted by strict rules, and so should "big data" innovators, to limit power imbalance.

There is a vital societal interest in efficient regulation of all new technologies, especially in the sphere of medicine. Many tech companies are either multinational corporations, or part of such, and pose political challenges for policy-makers as their regulation requires international cooperation and strong political will. Furthermore, innovative data processing and data mining products do not fit into existing definitions in legislation due to their specific features, and are raising questions as to, *e.g.* liability for AI. Doctors who base their diagnosis on AI processing of EHRs, medical journals, *etc.* may no longer be entirely liable for medical errors, as the fault may lie with the software development company. It is clear that the latter would prefer to self-regulate and deny having control over the data processing, but the unambiguous allocation of liability in the legislation can prevent "opening the floodgates" of costly litigation between the affected parties.

---

[37] WP29 nr 131 (N12), 5.
[38] Jason Chung, 'What Should We Do About Artificial Intelligence in Health Care?' (January 30, 2018). NYSBA Health Law Journal, Winter 2017, Vol. 22, No. 3. <https://ssrn.com/abstract=3113655> Last accessed on 18 March 2018.

Turning to the secondary-purpose processing, various stakeholders can seek to benefit or be adversely affected, depending on the purpose, and the type of information contained in the specific EHR dataset. As the Article 29 Data Protection Working Party has already observed, health information in EHRs "might generally attract the interest of third parties such as insurance companies and law enforcement agencies".[39] Although there is a public interest in resolving insurance fraud disputes more efficiently, for example, and insurers can benefit from disclosure of all insured's EHRs, many of the insured can object to insurance premiums set on the basis of advanced analytics of their personal health information. The application of "big data" processing to EHRs can also significantly disrupt the pharmacological sector. In contrast to controlled drug trials, analysis of medical histories, archive of prescriptions and treatment outcomes, over much longer periods of time, can offer granular data, enhance drug safety and change or expedite approval procedures.[40] This is clearly in the patients' best interests, but on the other hand, "big pharma" companies could also face more litigation when analysis confirms adverse drug effects.

These are just a few examples of the complex intermingling of public and private interests at stake. The conflicting motivations of the participants in 'big data' processing of EHRs projects, and certain third parties, increase the probability of utilisation of health data in ways that are incompatible with the initial context in which they were collected. That is why efficient regulation which addresses the issues arising out of such situations is essential for the healthcare domain.

2.2. CONTEXT TRANSGRESSION

As mentioned above, Nissenbaum focuses on context integrity in her framework of privacy. Thus, the right to privacy can be viewed as a right to flows of personal information appropriate within a specific context, the key constructs of which are roles, activities, norms, and values (sometimes called purposes).[41] When the contextual norms which govern such flows between

[39] Art 29 WP 131 (N30), 5.
[40] See, e.g. N Szlezák, M Evers, J Wang, L Pérez, 'The Role of Big Data and Advanced Analytics in Drug Discovery, Development, and Commercialization', (2014) Clinical Pharmacology & Therapeutics; 95 5, 492–495; B Chen, AJ Butte, 'Leveraging big data to transform target selection and drug discovery', (2016) Clinical Pharmacology & Therapeutics, 99 3, 285-297.
[41] Nissenbaum (N22), 133.

the sender and recipient are contravened, there is a transgression (violation) of the context, i.e. a violation of privacy.[42]

The relationships and roles of the actors (sender, recipient and data subject) are important in establishing context transgression.[43] Individuals rarely expect their medical records to be shared with anyone outside the doctor-patient relationship,[44] so when the opposite is true it raises the assumption that their privacy has been violated, especially when the recipients do not fulfil the duties of a physician. The utilisation of digital patient records seems to have made the latter inevitable.

"[…] EHR systems additionally have the potential not only to process more personal data (*e.g.* in new contexts, or through aggregation) but also to make a patient's data more readily available to a wider circle of recipients than before."[45]

Indeed, a growing number of parties will seek to benefit from the advanced data analysis of EHRs and its disclosure, in contravention to the norms regulating the informational flows of the data in the personal health record. Nevertheless, the ethical requirement of medical confidentiality, originally set out in the "Hippocratic Oath", has been maintained even in the newly revised World Medical Association's Declaration of Geneva, released in October 2017.[46] It excludes all third parties unless there is a specific legal basis or the patient's consent for the use of the information.[47] Still, in many of the recent health data initiatives, the information in EHRs is processed for secondary purposes such as scientific research or medical device/software development. The informational norms and legislation governing traditional medical research are in conflict with those regulating patient records; those regarding

---

[42] Nissenbaum (N22), 127.
[43] "[…] businessman to employee, minister to congregant, doctor to patient, husband to wife, parent to child, and so on. In each case, the sort of relationship that people have to one another involves a conception of how it is appropriate for them to behave with each other, and what is more, a conception of the kind and degree of knowledge concerning one another which it is appropriate for them to have." James Rachels,"Why Privacy Is Important." Philosophy & Public Affairs Vol. 4, No. 4 (Summer, 1975), pp. 323-333; 328.
[44] Entzeridoua (N29).
[45] Art 29 WP 131 (N29), 5.
[46] "[…] I will respect the secrets that are confided in me, even after the patient has died […]".World Medical Association's Declaration of Geneva, as amended by the 68th WMA General Assembly, Chicago, United States, October 2017. < https://www.wma.net/policies-post/wma-declaration-of-geneva/ > Last accessed on 20 March 2018.
[47] Art 29 WP 131 (N29), 10.

commercial software development are even more so. Furthermore, the digitalisation of medical records was the first step towards interoperability (shareability), which has become the focus of EU's eHealth agenda and will contribute towards new cross-border informational flows between providers.

Each of the participating organisations in "big data for healthcare" projects sets its own targets to be achieved in order to protect or promote its own values. On the one hand, the healthcare systems and organisations are strictly regulated in line with the public health policy and goals set by the national (and sometimes regional) legislative bodies, but on the other hand, commercial research and development projects follow profit-making imperatives and private agendas. "Hybrid" partnerships between public and private actors are therefore ridden with contradictory long-term strategies from their very inception. Having highlighted the conflicting values and goals of the multiple stakeholders in the application of commercial data science in the health domain, the likelihood of context transgressions seems high.

Turning to the activities, it is evident that when data mining, AI and machine learning are applied in processing of EHRs, these also differ from the procedures and practices of treating medical staff. The analysis of aggregate data about a single patient as part of a big dataset (consisting of potentially thousands, even millions of patient records) can be much more invasive than review of the patient file by a physician. Unlike a doctor, AI can easily process multiple linked "big" datasets containing hundreds of medical journals, EHRs, e-prescriptions, patient's identifiers *etc*. and reveal patterns that are impossible to detect otherwise.

Context transgression in the case of processing of EHRs can have a number of undesirable consequences for all stakeholders in the healthcare domain. One of the regulatory tools which seek to prevent negative impact in such cases is data protection law, which will be discussed in the following chapter.

2.3. CONTEXTUAL INTEGRITY AND EU REGULATION

It should also be stressed how the contextual integrity framework developed by the US ethicist Nissenbaum, which is already highly influential in the USA,[48] relates to data protection and privacy regulation in the European Union and the United Kingdom.

On the one hand, the theory of contextual integrity offers a nuanced approach to the prioritizing and balancing of fundamental rights and freedoms of individuals, public and private commercial interests. As has been observed in a report of a Brussels-based think-tank on European policy, its "flexibility […] resonates very well with stakeholders".[49] Moreover, examining legislation in light of the social norms and expectations in the specific context, one can better assess the effectiveness of that hard law. Further, Nissenbaum herself has suggested that the contextual integrity framework is intended to be used as a standard against which to evaluate legislation regulating data flows.[50]

Secondly, as has been noted,[51] the CJEU already tends to refer to 'all the circumstances'[52] of a case in its reasoning. It has been argued that the principle of respect for context was firmly embedded in the DPD.[53] It is even more obvious that the GDPR contains references to context and data subjects' expectations of data flows in a number of Recitals and Articles, in particular in Recital 38 pertaining to Article 6(1)(f) (which is the equivalent to Article 7(f) of the DPD).[54] This is not at all a coincidence. There is a discourse and exchange of ideas between US and EU

---

[48] Note, inter alia, the 'respect for context' clause in the February 2012 Consumer Privacy Bill of Rights Act proposal by the US Government. The White House, Consumer Data Privacy in a Networked World: a Framework for protecting privacy and promoting innovation in the global digital economy, February 2012, < https://www.whitehouse.gov/sites/default/files/privacy-final.pdf>. President Obama revived the proposal in March 2015, but the draft never became legislation and is now presented as a Framework, see at <https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/images/Documents/Privacy%20in%20Our%20Digital%20Lives.pdf> Last aaccessed on 10 May 2018.

[49] K. Irion and G. Luchetta, 'Online personal data processing and EU data protection reform' (8 April 2013) CEPS Task Force. Report of the CEPS Digital Forum. Centre for European Policy Studies 2013, at 57, <http://ssrn.com/abstract=2275267> Last accessed on 11 May 2018.

[50] H. Nissenbaum, *Privacy in context: Technology, policy, and the integrity of social life*, (Stanford University Press, 2009) at 16 and at 236.

[51] See Audrey Guinchard, 'Contextual Integrity and EU Data Protection Law: Towards a More Informed and Transparent Analysis' (6 March 2017). <https://ssrn.com/abstract=2946772> Last accessed on 10 May 2018.

[52] For example, Rechnungshof (C-465/00) v Österreichischer Rundfunk and Others, paras. 67, 76; Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González (C-131/12), para 94.

[53] K. Irion and G. Luchetta, (N49) at 57.

[54] Also noted in A. Guinchard (N51) at 13.

academics, as is evident from the recent interest in the contextual integrity theory by EU scholars.[55] The same is true of policy-makers.[56]

The context integrity framework can therefore be deemed to be a good tool for assessing the efficiency and limitations of data protection regulation. The discussion of some of the potential consequences of context transgression in the next chapter will highlight the ripple effects on society and the magnitude of the problems. Answering the main research question thus entails two tasks, the first of which is the legal analysis of the privacy and data protection framework in terms of efficient prevention of context transgression consequences. The second task involves the assessment of the legal response to context transgression impact once it has materialized. Each of these will be examined in turn in the fourth chapter.

[55] See, e.g. Sharon (N5), at page 6; Guinchard (N51).
[56] Cited in A.Guinchard (N51) In the discussion before the EU Parliament on 9 and 10 October 2012, prior thus to Vice-President of the EU Parliament A. Alvaro's December 2012 amendment of Article 5 for a 'Respect for context' clause, the US Privacy Bill has been discussed with Alvaro present. See the minutes of the meetings on pages 5 and 6 <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=%2F%2FEP%2F%2FNONSGML%2BCOMPARL%2BPE-504.214%2B01%2BDOC%2BPDF%2BV0%2F%2FEN> Last accessed on 20 March 2018.

CHAPTER 3: IMPACT OF CONTEXT TRANSGRESSION RESULTING FROM COMMERCIAL "BIG DATA" PROCESSING OF EHRs BY TECH COMPANIES

3.1. IMPACT ON INDIVIDUALS

As outlined above, in the traditional context of healthcare, patients have a certain set of reasonable expectations about the parties who are going to view their medical record or sections of it (doctors, nurses, lab assistants, pharmacists *etc.*), about the ethical principles they are bound by, and the actions and purposes they are going to use the information for (diagnosis, lab tests, dispensing medicine, *etc.*). The "big data" processing by private companies in partnerships with healthcare providers is in stark contrast to these privacy expectations and this has several major implications.

The average person's reactions to context transgression can be "indignation, protest, discomfort, and resistance to technology-based information systems and practices".[57] For example, in 2014-2015 NHS England faced public and media backlash over the care.data programme,[58] which aimed to link patient information from all NHS providers, so it could be used for purposes beyond direct care.[59] The citizens' concerns signalled that it was unacceptable to most people for their health information to be shared with actors outside the direct treatment relationship, or for medical research, especially when this was done without explicit patient consent. The National Institute for Health and Care Excellence (NICE) Citizens Council, a non-departmental UK public body which provides a platform for expressing individuals' opinions, later confirmed that the main reasons for refusal to share their patient record were people's concerns about the future uses, as well as the possibility that it may be passed on or sold to other organisations.[60] There was further media outrage over the failure to fulfil patients' requests to opt out of sharing due to fears that their sensitive data might end up being used by

[57] Nissenbaum (N29), 140.
[58] See, e.g. Nick Triggle, 'Care.data: How did it go so wrong?' *BBC news*, (19 February 2014), <http://www.bbc.com/news/health-26259101> Last accessed on 25 March 2018.
[59] NHS England. The care.data programme. <https://www.england.nhs.uk/ourwork/tsd/care-data> Last accessed on 25 March 2018.
[60] National Institute for Health and Care Excellence (NICE) Citizens Council, 'What Ethical and Practical Issues Need to Be Considered in the Use of Anonymised Information Derived from Personal Care Records as Part of the Evaluation of Treatments and Delivery of Care?', Citizens Council Reports No. 18, (11 November 2015), 32. <https://www.ncbi.nlm.nih.gov/books/NBK401705/> Last accessed on 28 March 2018.

insurers or pharmaceutical companies, signalling societal unrest when self-determination is not an option for patients.[61] NHS England responded to published articles, stating that patients will "continue to be asked for their explicit consent to view their SCR [Summary Care Record] by healthcare professionals, for the purpose of clinical care only".[62] Shortly thereafter, the details of the Google DeepMind/Royal Free NHS London agreement emerged, revealing that the contracting parties had relied on an inappropriate legal basis (the so-called "implied consent") to legitimise the processing of health data by the tech company for the development of the Streams app. The deal, as mentioned above, naturally invoked harsh criticism.

Patients' dismay is partially caused by the departure from the guiding ethical principles in healthcare that such incidents illustrate. There is a distinctly deantological perspective underpinning the rules on patient autonomy and control over sensitive health data, requirements of informed and explicit consent, duties of care and of confidentiality *etc*. Medical staff remain bound by national and international[63] ethical codes of practice which embody these guidelines and moral imperatives. In contrast, it has been argued that the rhetoric adopted by the tech corporations, various other private actors in the healthcare domain, and even the EU (e.g. Recital 157 GDPR) [64] is notably utilitarian, with emphasis on the instrumental role of personal data to medical innovation and promises of maximising the net benefits of its use. Individuals' reasonable expectations about their information shared in the context of a doctor-patient relationship are subverted in the case of EHR processing by private tech companies, and medical personnel are faced with new challenges to their traditional role as guardians and recorders of all health data and insights in the patient files.

---

[61] Randeep Ramesh, 'NHS disregards patient requests to opt out of sharing medical records', (*The Guardian*, 22 January 2015) <https://www.theguardian.com/society/2015/jan/22/nhs-disregards-patients-requests-sharing-medical-records> Last accessed on 3 March 2018.

[62] NHS England, Statement from NHS England and the Health and Social Care Information Centre in response to the Daily Telegraph article, 'Tesco can see your medical records' (10 August, 2015), <https://www.england.nhs.uk/2015/08/response-dt-article/> Last accessed on 25 March 2018.

[63] See e.g. WMA Declaration of Geneva (N38).

[64] See Nadezhda Purtova, 'Health Data for Common Good: Defining the Boundaries and Social Dilemmas of Data Commons' (9 July 2016). in Ronald Leenes, Nadezhda Purtova, Samantha Adams (eds.) (2017) Under Observation - The Interplay Between eHealth and Surveillance, Springer ; Tilburg Law School Research Paper No. 15/2016. < https://ssrn.com/abstract=2807455> Last accessed on 10 June 2018.

Respect for personal autonomy is still one of the main principles in healthcare, and the underlying reason for *e.g.* requiring informed patient consent[65] for medical procedures. However, seeking informed and explicit consent for sharing of digital medical records not only requires a relatively high degree of specificity to satisfy GDPR's requirements (as will be explained in the next chapter), but has also become a controversial issue, because it can impose an administrative and financial burden, participation bias and thus impede large studies.[66] It is therefore not surprising that healthcare providers look for alternative legal bases for the sharing or processing of EHRs. Patients, in effect, might then be coerced into sharing their health data for second use processing and non-treatment purposes. Violating the right to an informed patient choice is detrimental in itself. However, a further problem with such coercion, even for goals which are *prima facie* in the public interest, is that the participation of actors with conflicting values raises questions as to the private beneficiaries' hidden intentions and gains. The lack of transparency and long-term profit-making agendas that tech corporations are notorious for, in addition to constant efforts to diversify their business activities, have made many of them seem unpredictable, omnipresent and omnipotent, hence untrustworthy. When they actively initiate processing activities of EHRs to pursue such agendas, they would be liable as data controllers for compliance with the general data protection principles in Article 5(1) GDPR (which, *inter alia*, seek to enhance transparency and protect the confidentiality and security of the data).

As a result of controversial projects in various countries, there has been a loss of confidence in both the involved healthcare institutions and the partnering organisations. Surveys have found that individuals in the UK generally trust public institutions more than they do private actors, especially when the latter have business interests in sectors outside medicine.[67] But collaborative projects have further exacerbated concerns that the confidentiality of patient

---

[65] Informed consent is traditionally associated with the Anglo-Saxon common law tradition, whereas data protection instruments in continental Europe typically grant individuals the rights to view, correct, control or delete their personal data. See Linnet Taylor, Luciano Floridi, Bart van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies*, (Philosophical Studies Series, Springer, 2017) 6.

[66] A. Docherty, 'Big Data – Ethical Perspectives', (2014) Anaesthesia, 69, 387–398, 390 <https://onlinelibrary.wiley.com/doi/pdf/10.1111/anae.12656 > Last accessed on 20 March 2018.

[67] Ipsos Mori Research for The Royal Statistics Society. "New research finds data trust deficit with lessons for policymakers." Ipsos MORI. < https://www.ipsos.com/ipsos-mori/en-uk/new-research-finds-data-trust-deficit-lessons-policymakers> Last accessed on 23 March 2017.

records is threatened, whether by cybersecurity risks or health data initiatives for authorised second use.[68]

In addition to discomfort and protest, deeper anxieties may also affect individuals whose records are analysed digitally. After all, advanced computational methods have been developing at such a rapid pace, and knowledge about their mechanisms and potential application has only just started entering the lexicon of the average person. Not so long ago, "Artificial Intelligence" was widely understood to mean only android robots from sci-fi literature and films. Additionally, machine learning products are also often called "black box technologies", because they are so opaque even their developer may not comprehend the logic behind the resulting analysis.[69] It is therefore very difficult for lay people to truly appreciate and weigh the potential benefits or risks that advanced processing of their EHRs may entail.

As a consequence of the nature of the technologies applied, individuals may experience fears of both realistic and unrealistic threats from the context transgression. Cognitive biases can push people into panic just as much as into passivity.[70] Either one is undesirable and can impede successful and ethical implementation of EHR processing projects. The feeling of being monitored or the suspicion that their medical records will be sold to third parties can make some people circumspect[71] or even prevent them from disclosing symptoms or going to the doctor. This could have a detrimental effect on a person's health, and will likely affect certain minorities and vulnerable groups, *e.g.* patients with schizophrenia, who often have symptoms like being suspicious or withdrawn.[72]

---

[68] 'Survey sheds light on UK's levels of confidence in secure protection of health data', British Journal of Healthcare Computing, (22 May 2017) < http://www.bj-hc.co.uk/publics-trust-protection-health-records-reaches-worrying-levels> Last accessed on 10 August 2018.

[69] Neil Mehta, Murthy V.Devarakonda, 'Machine Learning, Natural Language Programming, and Electronic Health Records: the next step in the Artificial Intelligence Journey?', (5 March 2018), Journal of Allergy and Clinical Immunology, 2, <https://doi.org/10.1016/j.jaci.2018.02.025> Last accessed 25 March 2018.

[70] Cass R. Sunstein and Richard Zeckhauser "Overreaction to Fearsome Risks", (2008) HKS Faculty Research Working Paper Series, 1.

[71] Helen Nissenbaum, 'A Contextual Approach to Privacy Online' (2011). Daedalus 140 (4), Fall 2011: 32-48, 45 <https://ssrn.com/abstract=2567042> Last accessed on 25 March 2018.

[72] 'FDA approves pill with sensor that digitally tracks if patients have ingested their medication: New tool for patients taking Abilify' USA Food and Drug Administration website, (13 November 2017) <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm584933.htm> Last accessed on 29 April 2018.

One of the "nightmare scenarios" that trouble individuals, is for their sensitive health data to be digitally analysed by their respective employers.[73] Turning back to e-commerce giant Amazon's recent efforts, it is in the process of turning that fear into reality for its US employees. It was announced on 30 January 2018 that Amazon partnered with Berkshire Hathaway and J.P. Morgan Chase to reduce the costs of healthcare for employees of the three companies in the United States.[74] Although the initial press release stated that the new joint venture would be an "independent company that is free from profit-making incentives and constraints", which would, however, be jointly led by executives from all three corporations, critics immediately pointed out that Amazon itself "went through the first 23 quarters of its existence operating free from profit-making incentives", [75] but eventually went on to become the second most valuable corporation in the world.[76] The motives behind these new healthcare innovation projects clearly remain reducing costs and generating a return of the investment in the future. Inside sources foreshadowed[77] the subsequent acquisition of the e-pharmacy PillPack[78] by Amazon, and predicted that plans for a new health insurance company may also be announced in the near future.[79] As for the processing of health records, Amazon, JPMC and BH (through the new joint

[73] Docherty (N47) 390; See also Oliver Ritchie, Sophie Reid and Lucy Smith (2015) 'Review of public and professional attitudes towards confidentiality of healthcare data', 31. < http://www.gmc-uk.org/Review_of_Public_and_Professional_attitudes_towards_confidentiality_of_Healthcare_data.pdf_62449249.pdf> Last accessed on 1 April 2018.
[74] Fortune Editors and Reuters, 'Amazon, Berkshire Hathaway and J.P. Morgan Are Forming a Non-Profit Health Care Venture' (*Fortune*, 30 January 2018), <http://fortune.com/2018/01/30/amazon-berkshire-hathaway-jpmorgan-nonprofit-healthcare/> Last accessed on 4 February 2018.
[75] Clifton Leaf, 'Amazon–JPMorgan–Berkshire Hathaway: What Their New Health Venture Really Means' (*Fortune*, 31 January 2018) <http://fortune.com/2018/01/31/amazon-jpmorgan-berskshire-healthcare/> Last accessed on 1 August 2018.
[76] Natasha Bach, 'First Microsoft, Now Alphabet. Amazon Passes Another Giant to Become The Second Most Valuable U.S. Company', (*Fortune*, 21 March 2018) < http://fortune.com/2018/03/21/amazon-second-most-valuable-company-after-apple/> Last accessed on 30 March 2018.
[77] See Samantha Liss, 'Amazon gains wholesale pharmacy licenses in multiple states', (*St. Louis Post-Dispatch*, 27 October 2017), < http://www.stltoday.com/business/local/amazon-gains-wholesale-pharmacy-licenses-in-multiple-states/article_4e77a39f-e644-5c22-b5e6-e613a9ed2512.html> Last accessed on 25 March 2018.
[78] Bruce Japsen, 'It's Official: Amazon Enters Pharmacy Business With PillPack Acquisition', (*Forbes*, 28 June 2018), < https://www.forbes.com/sites/brucejapsen/2018/06/28/its-official-amazon-enters-pharmacy-business-with-pillpack-deal/#6035c07113fa> Last accessed on 1 July 2018.
[79] Fortune Editors and Reuters, 'Amazon, Berkshire Hathaway and J.P. Morgan Are Forming a Non-Profit Health Care Venture' (*Fortune*, 30 January 2018), <http://fortune.com/2018/01/30/amazon-berkshire-hathaway-jpmorgan-nonprofit-healthcare/> Last accessed on 4 February 2018.

venture) will likely store the EHRs[80] of all their employees and their employees' families, sell and deliver their prescribed medicine, as well as use Amazon's machine learning and "big data" analytics expertise to conduct medical research and develop innovative health technologies, *etc*. As the primary medical privacy legislation in the USA, the so-called HIPAA,[81] is only binding on those entities explicitly mentioned in the Act but not to other entities, *e.g.* online health services,[82] the joint venture's processing activities may not even be covered by HIPAA.

Given the porous boundaries among subsidiaries of large diverse companies like Amazon and Google, there is typically little to stop the free flow of information,[83] and the aggregation of data from multiple sectors and contexts. On the one hand, personal data is generally considered more valuable for commercial purposes. For instance, certain health and genetic data can be associated with various personality traits and moods like irritability, depression and stress *etc*.[84], so it could be used to predict an employee's future performance and behaviour (thus contributing to stigmatisation and potentially adverse material consequences for already vulnerable individuals). On the other hand, even the transfer of an anonymized dataset of digital health records to subsidiaries for secondary processing can be highly profitable[85] for Amazon and its partners, and the e-commerce giant is known for extracting value from data analytics. For example, identifying unknown drug adverse events through machine learning can provide a great competitive advantage to Amazon's pharmaceutical branch. Linking and cross-

[80] Eugene Kim and Christina Farr, 'Amazon has a secret health tech team called 1492 working on medical records, virtual doc visits', CNBC <https://www.cnbc.com/2017/07/26/amazon-1492-secret-health-tech-project.html> Last accessed on 2 February 2018.
[81] Health Insurance Portability and Accountability Act (HIPAA).
[82] Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013) at 34-35.
[83] Nissenbaum, Helen, "A Contextual Approach to Privacy Online" (2011). Daedalus 140 (4), Fall 2011: 32-48, 44. <https://ssrn.com/abstract=2567042> Last accessed on 25 March 2018.
[84] Craig Konnoth, 'Health Information Equity', 165 U. Pa. L. Rev. 1317 (2017), 1343 <http://scholar.law.colorado.edu/articles/701> Last accessed 2 February 2018.
[85] See, *e.g.* Kristin Lacy-Jones, Philip Hayward, Steve Andrews, Ian Gledhill, Mark McAllister, Bertil Abrahamsson, Amin Rostami-Hodjegana, Xavier Pepine, 'Biopharmaceutics data management system for anonymised data sharing and curation: First application with orbito IMI project', (2017) Computer Methods and Programs in Biomedicine, Volume 140, March 2017, Pages 29-44; Khaled El Emam and Sam Rodgers, 'Anonymising and sharing individual patient data', (2015) The BMJ, < https://www.bmj.com/content/350/bmj.h1139.abstract> Last accessed on 1 June 2018; National Institute for Health and Care Excellence (NICE) Citizens Council, "What Ethical and Practical Issues Need to Be Considered in the Use of Anonymised Information Derived from Personal Care Records as Part of the Evaluation of Treatments and Delivery of Care?", Citizens Council Reports No. 18, (11 November 2015), 32. <https://www.ncbi.nlm.nih.gov/books/NBK401705/> Last accessed on 28 March 2018.

referencing that dataset with non-sensitive information about the same data subjects (the so-called linkage attack)[86] can in effect reverse anonymization, or result in what has been called "reachability" (the possibility of holding you accountable or impacting you with or without access to identifiable information).[87] In short, if Amazon analyses the EHRs of its employees, in addition to information obtained through the employment relationship, consumer services, online pharmacy *etc.* and does so mainly with the intention of improving efficiency and reducing costs, then the potential risks to the data subjects extend far beyond serious privacy violations and harms.

Turning to the EU and UK legal frameworks, if Amazon wanted to do the same, in light of the power imbalance of the employer-employee relationship, consent to processing by the joint venture is still unlikely to be considered free and voluntary enough to be a valid legal ground under either regime. The UKDPA expressly[88] prohibits attempts to coerce (potential) employees or contractors into providing health records, so if the processor is an associated joint venture instead of the direct employer there may not be a significant distinction in practice, and it may not have a legal basis to process the records unless it sought similar partnerships to those between DeepMind/NHS. On the other hand, as discussed in the next chapter, Article 9(2)(h) GDPR, and Section 2(2)(b) of Part 1 of Schedule 1 UKDPA, respectively, may be broad enough to legitimise, *e.g.* secondary-purpose data analytics for the assessment of the working capacity of employees. Furthermore, Article 9(2)(j) GDPR may allow analysing health data for scientific or statistical purposes, subject to the requirement for data minimisation (pseudonymisation or anonymisation). Some people would simply consider the digital comprehensive record of their health status and care (which may in future include DNA sequencing data) to be far more sensitive personal data than their name, residence address *etc.*, so they would question the possibility of anonymising a uniquely identifying set of details. Their negative reaction to context transgression cannot be avoided.

[86] Further discussed in Solon Barocas, Helen Nissenbaum, "Big Data's End Run around Anonymity and Consent" pp.44-75 in Julia Lane et.al. "Privacy, Big Data, and the Public Good: Frameworks for Engagement", (Cambridge University Press 2014).
[87] *Ibid*, 51.
[88] Section 184(1) UKDPA and Section 185 UKDPA, discussed in the next chapter.

The actual impact of EHR processing by tech companies depends on a number of country-specific factors, including the types of data added to the electronic health record (whether it contains *e.g.* data from wearable devices such as FitBit, or lifestyle details), the applicable laws regulating patient records, the exact terms of the data-sharing agreement, the attitudes towards privacy and sharing sensitive data prevalent in the country, *etc.*

3.2. IMPACT ON GROUPS OF INDIVIDUALS AND SOCIETY

In addition to the potential impact on each individual, there may be further consequences on the group level, and for society in general. For example, data mining with algorithms aimed at improving the cost efficiency of healthcare systems can contribute to marginalization of groups that require expensive treatment or are at higher-than-average risk of developing a certain condition. Thus, one of the potential detrimental effects of EHR analysis is stigmatisation of highly vulnerable groups of individuals, even though the goal of collection of health data is to improve their well-being.[89] Medical research conducted in Scotland by linking EHRs and census datasets discovered some years ago that the incidence of acute myocardial infarction is higher among South Asian residents.[90] The argument has been made that this creates a possibility to misuse such findings against the interests of the ethnic minority groups concerned.[91] It is however doubtful whether the GDPR and the UKDPA may legitimise EHR processing for such improper purposes on the basis of the "public interest in the area of public health" exception, in light of GDPR's stated aim[92] of protecting fundamental rights.

Whereas traditional self-aware groups, such as ethnicities, are widely recognised in society and may be protected by law, pattern discovery through machine learning can form groups that are

[89] A. Docherty, "Big Data – Ethical Perspectives", Anaesthesia 2014, 69, 387–398, <https://onlinelibrary.wiley.com/doi/pdf/10.1111/anae.12656 > Last accessed on 20 March 2018.
[90] Fischbacher et al, 'Record linked retrospective cohort study of 4.6 million people exploring ethnic variations in disease: myocardial infarction in South Asians', BMC Public Health 2007 < https://doi.org/10.1186/1471-2458-7-142> Last accessed on 20 March 2018.
[91] Kenneth M Boyd, 'Ethnicity and the ethics of data linkage', BMC Public Health 2007 7:318 <https://doi.org/10.1186/1471-2458-7-318> Last accessed on 20 March 2018.
[92] Article 1(2) GDPR.

imperceptible to humans (such as genetic categories)[93] and whose rights are not yet legally safeguarded. The application of advanced data analytics has changed the way groups can be defined and identified.[94] Further, when the identification of the group is only a preliminary stage for achieving a specific purpose, even the data scientist may be unaware of it,[95] not to mention the medical staff interpreting the final analysis or the patients themselves. As a consequence, any implicit biases "inherited" from the data source or unintentionally embedded in the algorithm will go unnoticed[96] by the data controllers, processors, and most importantly, by the group affected by the processing. If the purpose of the processing is medical research, these biases will be passed on in all diagnoses informed by the study, which in turn will be recorded in the EHRs and processed for various new purposes, and so on. Due to the effect of the feedback loop, it should be kept in mind that any biases and errors may be passed on in perpetuity, despite the accuracy principle in the GDPR.

Algorithm-driven profiling on the basis of "big data" has become a revolutionary tool for mass persuasion[97] and has been utilised by technology-centred companies in many sectors, from e-commerce to politics.[98] When such companies apply the same powerful computing methods to classify patients and their behaviour on the basis of data in the EHRs, naturally various stakeholders are interested in exploiting the group profiles. One recent example is the successful development of a predictive model (based on 300 different factors) that can help determine the likelihood that a patient will not fill a prescription. The factors were identified using an algorithm by the US-based company Express Scripts, which administers 1.4 billion prescriptions for 100 million patients per year, and which announced its sale to insurer Cigna

[93] See Dara Hallinan and Paul de Hert, *Ch. 10 Genetic Classes and Genetic Categories: Protecting Genetic Groups Through Data Protection Law* in Linnet Taylor, Luciano Floridi, Bart van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies*, (Philosophical Studies Series, Springer, 2017).
[94] Linnet Taylor, Luciano Floridi, Bart van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies,* (Philosophical Studies Series, Springer, 2017) 41.
[95] *Ibid* 42.
[96] Patricia Balthazar et al., 'Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics', (2018) Journal of the American College of Radiology, (March 2018 Volume 15, Issue 3, Part B, Pages 580–586) 584, < https://doi.org/10.1016/j.jacr.2017.11.035> Last accessed on 28 March 2018.
[97] Taylor (N94) 4; 10.
[98] 'Facebook under fire in escalating data row' (*BBC News*, 19 March 2018) <http://www.bbc.com/news/technology-43461865> Last accessed on 29 March 2018.

in March 2018.[99] The group of patients who fit into the high risk category are assigned scores and receive targeted reminders - "soft touch, nothing Orwellian", according to the chief data officer of the company. As a result of these paternalistic but efficiency-driven efforts, non-adherence to doctor's prescriptions have been reduced by 37%, the costs of treatment of medical complications due to non-compliance have also decreased. Additionally, the insurers partnering with Express Scripts have reported major savings.[100] Presumably, the same profiling algorithm can enable insurers to impose higher premiums or sanctions for repeat offenders who do not fill out their prescriptions, or enable pharmaceutical companies to choose suitable drug safety and clinical trial participants. Depending on the application of the profiling results, the data controllers and the sectors where the outcomes are used, the direct and indirect consequences for the patients can vary significantly. In this instance, there seems to be a direct positive impact on the health (or at least the habits) of the disobedient patients, however it is at the expense of the right to privacy of all the data subjects. The privacy violations can have a number of indirect consequences for patients, as discussed above and hereafter.

Whereas the application of AI for the purposes of improving compliance with prescription medicine may be legitimised on grounds of almost all of the GDPR exceptions to the sensitive data general prohibition (vital interests; preventive medicine; substantial public interest; public interest in the area of public health; statistical purposes), any non-anonymized EHR data processing by insurers or pharmaceutical companies may be inappropriate and likely be in conflict with the purpose limitation principle.

Conclusions about groups (on the basis of ethnicity, age, residence, lifestyle, *etc.*) resulting from "big data" research can inadvertently provoke prejudice, dignitary/reputation/status harm and discrimination towards all individuals who possess the shared characteristics. Even patients who have not consented to processing of their health record and whose EHR has not in fact been processed can be affected.[101] Although the rights to data protection and medical confidentiality of the latter may not be breached, they could still suffer adverse consequences

[99] Erika Fry and Sy Mukherjee, 'Tech's Next Big Wave: Big Data Meets Biology' (*Fortune,* 19 March 2018) <http://fortune.com/2018/03/19/big-data-digital-health-tech/> Last accessed on 28 March 2018.
[100] *Ibid.*
[101] For a discussion on "the network effect", which refers to the fact that the loss of privacy of one individual may have an impact on the privacy of others, see, *e.g.* Taylor (N94) 9.

if EHR processing inferences are utilised in inappropriate contexts, such as employment or credit rating services. As will be discussed in the next chapter, the UKDPA seems to allow second-purpose processing of EHRs in the employment context.

Turning to reactions to context transgression, certain groups are more likely to have a positive attitude towards data-driven initiatives than others. Millenials, for instance, are said to be more open to "donating" their data for research and various causes.[102] However, this is not true in all countries.[103] In the USA, for instance, the wealthier and younger groups are less likely to have their health data in the public data pool.[104] This has raised the issue of health information equity in the States – *i.e.* distributing the informational burden in a just manner, as opposed to exposing, in this case - the poor and the elderly, to a disproportionately high risk of negative impact, although the whole of society may benefit from innovative medical research.[105] Especially controversial is the fact that it is the wealthy groups that receive personalised medical treatments even before they become widely available as a result of the "trickle-down effect", yet they are least likely to be adversely affected. In Europe, health information inequities or attitudes are country-specific and outside the scope of data protection law.

It should be acknowledged that simply following the so-called reductionist interpretation of the value of privacy by examining the undesirable consequences of context transgression (such as the aforementioned personal distress and social injustice) has its limitations. [106] For example, it leaves outside its scope the appreciation of privacy as a natural right to have exclusive control (ownership) over one's own personal data.[107] Further still, it does not take into account the idea that each person or each group is constituted by his/her or its information, respectively, and that

---

[102] This rhetoric is adopted by tech companies, see *e.g.* the IBM Watson executive's statement that, "the generation who buy Apple Watches are interested in data philanthropy". D. Crow, "IBM strikes digital health deal with Apple, Medtronic and J&J.", *Financial Times*. <www.ft.com/cms/s/0>.
[103] See Richie (N73), 33.
[104] Craig Konnoth, 'Health Information Equity', 165 U. Pa. L. Rev. 1317 (2017), 1332 <http://scholar.law.colorado.edu/articles/701> Last accessed 2 February 2018.
[105] *Ibid*.
[106] Taylor (N94) 92-95.
[107] For a detailed analysis, see *e.g.* Nadezhda Purtova, 'Illusion of Personal Data as No One's Property' (October 29, 2013). Law, Innovation, and Technology, Volume 7, Issue 1, 2015. < https://ssrn.com/abstract=2346693>; Gianclaudio Malgieri, 'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution? (November 20, 2016). Journal of Internet Law, Vol. 20, n.5, November 2016. <https://ssrn.com/abstract=2916079>.

a privacy violation is, therefore, an attack upon the identity of said person or group.[108] However, the suggestion of property rights and control over the record is unsustainable in respect of EHRs subject to advanced data analytics by tech companies in collaboration with healthcare institutions. The identity-constituting conception of privacy, on the other hand, could not address the position of all the stakeholders in the specified context. These theories have therefore not been discussed in detail in this thesis.

3.3. IMPACT ON HEALTHCARE PROVIDERS AND INSTITUTIONS

Healthcare providers and institutions who partner with tech corporations for advanced analytics projects are affected in several significant ways.

To begin with, there are ramifications for the performance of the professional and ethical duties of doctors, as well as for the legal obligations of healthcare institutions. They are directly affected when deemed to be the data controller for the purposes of data protection regulation, especially by the obligations on data controllers to be able to demonstrate compliance with the GDPR.

As mentioned earlier, data-sharing agreements concluded by hospitals for the application of machine learning or AI to health records challenge the role of physicians as guardians of the information contained in patient files. General practitioners and their practices, for example, may still be expected to play the role of the patient's advocates,[109] even when they are not parties to health data-sharing agreements for research purposes. Further, the more strategic investment there is for such projects,[110] the greater the emphasis on the physicians' duties to maintain comprehensive records will be. However, doctors find using EHR systems too time-

---

[108] Taylor (N94) 94.

[109] N. Mathers, G. Watt, N. Perrin, 'Towards consensus for best practice: use of patient records from general practice for research', (Welcome Trust, 2009) <https://wellcome.ac.uk/sites/default/files/wtx055661_0.pdf> Last accessed on 1 April 2018. Patricia Balthazar et al., 'Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics', Journal of the American College of Radiology, March 2018, Volume 15, Issue 3, Part B, Pages 580–586, 582 <https://doi.org/10.1016/j.jacr.2017.11.035> Last accessed on 30 March 2018.

[110] J. Oderkirk, 'Readiness of electronic health record systems to contribute to national health information and research', (2017), OECD Health Working Papers, No. 99, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9e296bf3-en> Last accessed on 4 July 2018.

consuming[111] and when pressed for time, they prioritise fulfilling their therapeutic obligations to patients over keeping their notes complete.[112] This tendency is in conflict with the data accuracy principle in the GDPR and poses major obstacles to extracting valuable insights for medical research *etc.* because the accuracy of the secondary use processing depends to a great extent on the format and quality of the source data. Data protection regulation and "Big Data" projects thus shift the focus of the doctors' duties from medical confidentiality to administering records.

Medical staff may also express privacy concerns in reaction to context transgression. An interesting case cited by Nissenbaum brought up the question whether doctors' prescriptions could be considered their personal information.[113] The physicians' complaint to the Privacy Commissioner of Canada stressed that data analytics of prescriptions was carried out for the purpose of discerning prescription patterns and the results were sold to foreign pharmaceutical companies. One of the algorithms tracked monthly prescribing activities of physicians who had attended events sponsored by participating pharmaceutical companies.[114] Although the Privacy Commissioner rejected the doctors' complaint on the basis that "personal information" was only information "about" a person, not information merely associated with a person, an alternative analysis could focus on context transgression and the right to group privacy instead. It should be noted that the physicians were all ranked and divided into groups depending on the average number of prescriptions they had written for drugs in a specific therapeutic class.[115] Even if the Commissioner's claim that insight into prescription habits relates to "work product" is correct, and the information fails the identity-defining test in Canada (or similarly does not fall under the Article 4(1) GDPR definition in the EU and UK), it can still be argued that selling it to a foreign pharmaceutical company constitutes a context transgression. The impact on the

---

[111] Erika Fry and Sy Mukherjee, "Tech's Next Big Wave: Big Data Meets Biology" (Fortune, 19 March 2018) <http://fortune.com/2018/03/19/big-data-digital-health-tech/> Last accessed on 28 March 2018.
[112] Ian P. McLoughlin, Karin Garrety, and Rob Wilson, *The Digitalization of Healthcare Electronic Records and the Disruption of Moral Orders* (Oxford University Press, 2017) 148.
[113] Nissenbaum (N29) 156-157.
[114] *Ibid.*
[115] Office of the Privacy Commissioner of Canada, "Privacy Commissioner releases his finding on the prescribing patterns of doctors", PIPEDA Case Summary #2001-15, Ottawa, October 2, 2001 <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2001/wn_011002/> Last accessed on 30 March 2018.

doctors can be, *inter alia*, a course-correction, *e.g.* more frequently prescribing drugs, even if unnecessarily, so as to earn favour with the pharmaceuticals. Such an adverse effect may be outside the scope of the GDPR and UKDPA. On the other hand, if the same analysis is utilised by the hospitals employing the physicians for the purposes of detecting unsafe practices and medical negligence, then the purposes and information flow would be context-appropriate.

Healthcare institutions are typically expected to maintain transparency,[116] to uphold patients' rights[117] and will ultimately be held accountable for the health record security under the GDPR and UKDPA. However, as highlighted above, tech titans are especially secretive when it comes to their long-term agendas, and they specialise in "black box technologies", making it practically impossible for the institutions to ensure transparency about all the (potential) applications of the analytics or oversight of the purposes of the processing. This is problematic not only for patients and society, but also for the healthcare institutions. The lack of technological expertise may place the latter in a subordinate position in relation to the "big data" companies, even in situations when they are considered to be the data controllers and are liable as such.

3.4. IMPACT ON THE TECH COMPANIES

Tech companies carrying out the processing of EHRs stand the most to benefit from the data analytics. To begin with, using clinical data in digital records, which is collected, assessed and input by medical experts, contributes to the validation of the subsequent analytics in the eyes of the scientific community. When the processing involves partnering with practicing clinicians to develop algorithms further adds credibility to the resulting analysis and brings a multi-disciplinary approach. The EHR datasets, the algorithms, any devices and products developed for secondary use processing can be extremely valuable, so tech companies will seek to retain control or intellectual property rights over their exploitation.

---

[116] L. Dauwerse, T. A. Abma, B. Molewijk, G. Widdershoven, 'Goals of Clinical Ethics Support: Perceptions of Dutch Healthcare Institutions', (2013) Health Care Analysis, December 2013, Volume 21, Issue 4, pp 323–337, <https://link.springer.com/article/10.1007/s10728-011-0189-5> Last accessed on 30 March 2018.

[117] Mathers N, Watt G, Perrin N "Towards consensus for best practice: use of patient records from general practice for research", (Welcome Trust, 2009) <https://wellcome.ac.uk/sites/default/files/wtx055661_0.pdf> Last accessed on 1 April 2018.

The existing legal frameworks may not adequately allocate or protect those rights, or there may be legal uncertainty making investment in development and participation in data processing projects riskier. The GDPR and the UKDPA, respectively, have brought a certain degree of novelty (hence, some legal uncertainty), so tech companies are bound to experience some difficulty in complying with the new obligations of data controllers and data processors.

The storage limitation principle in the GDPR may be difficult to comply with in the case of EHRs, when each medical fact may be relevant for a specific time period, or indefinitely. In any case, with so many stakeholders interested in EHR analytics, the companies processing the datasets are going to explore various ways to commodify their research. However, as the UKDPA introduced criminal offences in respect of unlawful obtaining of personal data, or retaining or selling on such data to third parties (discussed in the next chapter), employees of tech companies will think twice before attempting such conduct in the UK. Identifiable data may be more valuable, but under the UKDPA individuals attempting to reverse anonymisation or pseudonymisation of personal data will also be criminally liable for such conduct. There are also restrictions in the healthcare domain which may limit the opportunities to legally exploit all the tech companies' information assets.

As the profitability of EHR research and the value of datasets keep increasing and the projects undertaken by high-profile tech companies typically attract much public attention, the repositories storing medical records will also become the prime targets of various types of hacker attacks.[118] Apart from the "theft" and black market sale of information contained in EHRs, DDoS attacks, extortion, identity theft *etc.*, there is also the possibility of corrupting AI-driven research. According to tech experts, AI is "easy to fool",[119] which makes processing vulnerable and the analytics – corruptible. Tech companies will, therefore, have to ensure the highest level of cyber security and data governance in light of the increased likelihood of

[118] See 'Are EHR Vendors Hackers' Next Big Target?', *Hit Consultant* (Apr. 11, 2016), <http://hitconsultant.net/2016/04/11/preparing-ehr-vendors-cyber-threats/> Last accessed on 30 March 2018.

[119] Edd Gent, 'AI Is Easy to Fool—Why That Needs to Change', (Singularity Hub, 10 October 2017), <https://singularityhub.com/2017/10/10/ai-is-easy-to-fool-why-that-needs-to-change/> Last accessed on 31 March 2018.

attacks, especially in cases when they are the data controllers or joint data controllers (and, as such, are liable to hefty fines for GDPR non-compliance, or damages claims *etc*.).

3.5. IMPACT ON THE PHARMACEUTICAL INDUSTRY

There can be numerous consequences of the application of advanced data analytics to health records for "big pharma", some of which have been touched upon hereto.[120] EHRs can be utilised for clinical trials, such as when testing new pharmaceutical products,[121] by providing much more comprehensive systematically collected information and improving drug safety studies. Furthermore, evaluating drug candidates against patient record and genetic datasets has already been proven to reduce R&D costs for medicine development.[122] In short, there are various potential benefits from the use of EHR datasets for the goals of the pharmaceutical industry. However, whether or not such use is legal or acceptable for patients depends on the country.

Under the GDPR, the new public interest in the area of public health[123] and scientific research[124] exceptions may offer a legal basis for processing by pharmaceutical companies, provided that there are valid grounds (outside the scope of data protection law) for sharing the EHR datasets.

3.6. IMPACT ON PUBLIC HEALTH

To have a positive impact on public health, partnering companies and institutions involved in "big data" projects first need to maintain public trust.[125] This is not only due to the ethical principles in the healthcare systems, but also due to the reactions of society. For example, the care.data programme was eventually closed down precisely due to loss of public trust.[126]

---

[120] See, e.g. N Szlezák, M Evers, J Wang, L Pérez, 'The Role of Big Data and Advanced Analytics in Drug Discovery, Development, and Commercialization'', Clinical Pharmacology & Therapeutics (2014); 95 5, 492–495. doi:10.1038/clpt.2014.29; B Chen, AJ Butte, "Leveraging big data to transform target selection and drug discovery", Clinical Pharmacology & Therapeutics (2016), 99 3, 285-297, <https://doi.org/10.1002/cpt.318>.
[121] OECD paper (N86) 39.
[122] See Erika Fry and Sy Mukherjee, "Tech's Next Big Wave: Big Data Meets Biology" (Fortune, 19 March 2018) <http://fortune.com/2018/03/19/big-data-digital-health-tech/> Last accessed on 28 March 2018.
[123] Article 9(2)(i) GDPR.
[124] Article 9(2)(j) GDPR.
[125] Tjeerd-Pieter van Staa, Ben Goldacre, Iain Buchan, Liam Smeeth, 'Big health data: the need to earn public trust', (14 July 2016), BMJ : British Medical Journal; London Vol. 354.
[126] *Ibid.*

There are also technical and practical obstacles to realising the full potential of machine learning and data mining to improve public health, such as data quality challenges (incomplete or incorrect EHR data, inherent biases in the clinicians' notes or algorithms, population biases), increased likelihood of cyber-attacks[127] *etc.* Therefore, over-reliance on machine learning and AI-driven research is not going to lead to significant improvements in healthcare provision.

3.7. IMPACT ON HEALTHCARE POLICY AND REGULATION

Many countries, even most of those which have started implementing digital health record systems, have reported that they lack or have limited technical, financial and human resources to develop datasets.[128] Commercial entities, especially large multinational tech corporations, are often the only partners that can help national healthcare systems with the implementation of "big data" projects. Although public-private partnerships in healthcare are not a new phenomenon and have the potential for delivering long-term projects,[129] national policymakers are likely to let overly ambitious e-Health plans cloud their judgment and lead to "lock-in" situations and lax regulation (with many derogations from the baseline of protection established in the GDPR). Novel legal issues surrounding machine learning and advanced data analytics, such as liability for AI errors in healthcare,[130] are currently outside the scope of the UKDPA or the GDPR (which only impose a requirement for data to be kept accurate and up to date by the data controllers), but will require international cooperation and regulation to be efficiently resolved and avoid a "regulatory race to the bottom". There are risks and harms associated with secondary-purpose processing of EHRs which constitutes a context transgression, so they need to be assessed and a proper balance needs to be struck between the interests of all key stakeholders. National ethical and legal frameworks need to be put in place to supplement the GDPR for successful long-term harnessing of the power of EHR analytics.

---

[127] 'Eleven of 14 NHS health boards hit by ransomware cyber-attack' (*BBC News*, 12 May 2017), <http://www.bbc.com/news/uk-scotland-39896639> Last accessed on 1 April 2018.

[128] OECD paper (N86) 37.

[129] See e.g. Roehrich, Jens and Lewis, Michael and George, Gerard, 'Are Public-Private Partnerships a Healthy Option? A Systematic Literature Review' (2014) Social Science & Medicine, Vol. 113, pp. 110-119. <https://ssrn.com/abstract=2955093> Last accessed on 1 June 2018.

[130] See, e.g., Chung, Jason and Zink, Amanda, 'Hey Watson, Can I Sue You for Malpractice? Examining the Liability of Artificial Intelligence in Medicine' (23 November 2017). Forthcoming, Asia-Pacific Journal of Health Law, Policy and Ethics. <https://ssrn.com/abstract=3076576> Last accessed on 30 March 2018.

CHAPTER 4: HOW DOES DATA PROTECTION REGULATION ADDRESS CONTEXT TRANSGRESSION CONSEQUENCES AND HOW EFFICIENTLY?

4.1. THE EU AND THE UK LEGAL FRAMEWORKS APPLICABLE TO EHRs

In Europe, digital medical records are regulated through healthcare laws, legislation on patient rights, data protection rules and general rules on privacy protection.[131] The Article 29 Working Party (the old EU data protection expert advisory body) noted in its opinion from 2007 on EHRs[132] that the EU legal framework applicable to them consisted of the general provisions relating to the right to personal data protection in Article 8 of the EU Charter of Fundamental Rights, (as well as the right to respect for private and family life, home and correspondence in Article 8 of the European Convention for the Protection of Human Rights, and the Council of Europe Convention 108[133] in its entirety), the specific rules in the EC Data Protection Directive 95/46/EC (the "DPD") and Directive 2002/58/EC on privacy and electronic communications (the "ePrivacy Directive"), in addition to the national laws of the Member States implementing these Directives. Following the extensive reform in the field of data protection, the DPD has been repealed and replaced by the GDPR, which came into force on 25 May 2018. Additionally, on 10 January 2017 the European Commission adopted a proposal for a Regulation on Privacy and Electronic Communications to replace the ePrivacy Directive[134] and bring it into line with the GDPR.

Although the GDP Regulation has direct effect in all Member States, additional rules and derogations may be implemented in national legislation. Turning to the United Kingdom, in light of Brexit and the parliamentary supremacy principle in the UK, an Act of the Westminster Parliament is also essential in order to supplement and ensure that the data protection rules remain applicable after UK's withdrawal from the European Union. The Data Protection Act

---

[131] Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013) at 27.
[132] Article 29 Working Party, Working Paper nr 131, "Working Document on the processing of personal data relating to health in electronic health records (EHR)", adopted on 15 February 2007, at 1.
[133] Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108) and the Additional protocol to Convention 108 regarding supervisory authorities and transborder data flows (ETS No. 181).
[134] Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications).

2018[135] ("UKDPA"), which received Royal Assent (became law) on 23 May 2018, [136] implemented the rules of the GDPR and the EU Law Enforcement Directive 2016/680.[137]

This chapter will focus on the most relevant provisions in the GDPR and the UKDPA which apply to the processing of electronic heath records. The comparative analysis will serve to illustrate their efficiency as safeguards and the ways that they address context transgression. Due to the fact that legislation can only serve its intended purpose of protection and balancing of rights if it is efficiently enforced, the institutions tasked with its interpretation, supervision and with dispute resolution should also be discussed in the analysis of data protection regulation.

4.2. INTERPRETATION AND ENFORCEMENT OF DATA PROTECTION LEGISLATION IN THE UK AND THE EU

The European Data Protection Board (EDPB)[138], which succeeded the Article 29 Working Party, is the decision-making EU body in charge of the application of the GDPR as of 25 May 2018. Its guidelines and opinions on the interpretation of the Regulation are binding.[139] It is made up of the head of each Member State's Data Protection Authority and of the European Data Protection Supervisor (EDPS) or their representatives. All members of the Board have the opportunity to contribute to and influence data protection rules and their interpretation throughout the Union, which is especially valuable in respect of efficient regulation of multinational tech companies. The UK's Information Commissioner currently has a seat in the EDPB and will continue to take part in the decision-making process up to the country's exit from the EU. As a third country post-Brexit, the UK would not have an official role to directly contribute to the EU regime, unless a special precedent deal is agreed. The extraterritorial

[135] Data Protection Act 2018 c.12 < http://www.legislation.gov.uk/ukpga/2018/12/contents/enacted/data.htm> Last accessed on 5 June 2018.
[136] UK Parliament website, Bill stages — Data Protection Bill [HL] 2017-19, <https://services.parliament.uk/Bills/2017-19/dataprotection/stages.html> Last accessed on 22 May 2018.
[137] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA.
[138] See Articles 63 to 76 and Recitals (135) to (140) of the GDPR.
[139] Recital 136 GDPR.

application of data protection principles is one of the advantages of supranational regulation such as the GDPR in respect of efficient context transgression safeguarding when multinational tech corporations process EHRs or transfer them across borders.

The CJEU remains the ultimate arbiter and the judicial authority[140] which will guide and ensure the uniform interpretation of the legal framework in all Member States. Its jurisdiction as the final adjudicator in disputes related to EU legislation, such as the GDPR, cannot be questioned. However, once the UK finalises its withdrawal from the legal order of the Union, it will no longer be under the jurisdiction of the CJEU. This has caused some concerns about the uniform interpretation of the data protection rules on both sides of the Channel in the long term. [141] The UK courts will no longer be able or be bound to submit preliminary ruling requests on the interpretation of the GDPR to the CJEU.[142] On the other hand, they will have discretion to offer their own views. As a consequence, the question arises as to who would bring an action against the UK if it infringed data protection rules by misapplying the GDPR.[143]

When the UK is no longer bound by the decisions of the chief judicial authority of the EU, then its own Supreme Court will, in theory, be free to establish new precedents in respect of data protection rules, even if they are in contradiction to the CJEU's decisions and interpretation of the Regulations. The possibility for divergence from the case law of the Court of Justice means that the UK Supreme Court may choose to prioritise differently the rights of stakeholders in cases concerning the processing of EHRs by tech companies in partnership with healthcare providers or institutions. To give an example, the purpose limitation principle (Article 5(1)(b) GDPR and Article 36 UKDPA) may leave some scope for the judicial determination whether the purpose of further processing is compatible with the initial legal basis or not. The Supreme Court can then give more leeway to the controllers by adopting a broad interpretation of compatibility, or restrict the secondary purposes through a narrow interpretation. The degree of

[140] See Recital 143 GDPR – Judicial Remedies.
[141] Speech by Michel Barnier at the 28th Congress of the International Federation for European Law (FIDE), Lisbon, 26 May 2018, European Commission Press Release <http://europa.eu/rapid/press-release_SPEECH-18-3962_en.htm> Last accessed on 3 June 2018.
[142] Recital 143 GDPR reiterates that when a case is brought before a national court (in a Member State) it may, or in certain cases must, request a preliminary ruling from the Court of Justice on the interpretation of EU law, including the GDPR.
[143] Speech by Michel Barnier (N141).

impact in cases of context transgression will then differ from Member States which adhere to another interpretation. The CJEU facilitates uniform interpretation and thus offer greater legal certainty throughout the Union compared to the UK.

Although the UK Data Protection Act ("UKDPA") is intended to keep the state in line with the reformed EU legislation, there would be implications for the transfers of personal data in EHRs from and to the EU if the UK does not obtain an adequacy decision[144] from the European Commission under Article 45 GDPR on or before Brexit Day, or at any point loses that status. Certain concerns in that regard were expressed over the course of the parliamentary debates on the draft bill.[145] Even before that, critics feared that certain gaps in the national legislation at the time and contradictions to the new EU regime would jeopardize the free flow of data between the UK and the EU post-Brexit.[146]

A further complication arises from Section 16 UKDPA, which gives the Secretary of State of the UK the power to make regulations altering the application of the GDPR including adding or varying the derogations in Schedules 2 to 4 and omitting provisions subsequently added by regulations. A privacy watchdog has been quick to criticise the broad delegated powers granted to the Secretary of State and bypassing parliamentary scrutiny.[147] Any exercise of such powers can potentially push the UK away from compliance with the EU data protection regime. The Secretary of State enjoys a very broad discretion under the UKDPA to add further exceptions to the general prohibitions of sensitive data processing. It is yet to be determined whether the

[144] "The European Commission has the power to determine, on the basis of article 45 of Regulation (EU) 2016/679 whether a country outside the EU offers an adequate level of data protection, whether by its domestic legislation or of the international commitments it has entered into.", European Commission website, Data Protection < https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en> Last accessed on 20 May 2018.
[145] See e.g. UK House of Commons Hansard, 22 March 2018, <https://hansard.parliament.uk/Commons/2018-03-22/debates/ACC7E864-F2E5-4766-8590-FF26CD6C4BB3/LeavingTheEUDataProtectionAgreements> Last accessed on 21 May 2018.
[146] Andrew D. Murray, 'Data transfers between the EU and UK post Brexit?', (2017) International Data Privacy Law, Volume 7, Issue 3, 1 August 2017, Pages 149–164, https://doi.org/10.1093/idpl/ipx015 Last accessed on 20 May 2018.
[147] Privacy International Makes Recommendations To Strengthen UK Data Protection Bill, 1 October 2017 <https://privacyinternational.org/press-release/626/privacy-international-makes-recommendations-strengthen-uk-data-protection-bill> Last accessed on 4 June 2018.

final version of the Data Protection Act will meet the EU Commission's approval, or if amendments will be required before an adequacy decision is granted.

The absence of an adequacy decision would make it more costly and more difficult[148] for tech companies to legitimise transfers of big data sets of EHRs from a Member State to a UK-based processor. Watson for Oncology, the cloud-based artificial intelligence platform developed by IBM to assist doctors by analysing, *inter alia*, patient records, has been introduced in Dutch hospitals,[149] and transfers of personal health records to and from a UK-based cloud server are not out of the realm of possibilities (considering that IBM has subsidiaries in both countries). Explicit informed consent of the patients may offer a potential legal basis for the transfers under Article 49(1) GDPR, as well as prevent adverse psychological reactions of patients or loss of public trust. It can, however, reduce the number of data subjects, and impose an administrative burden (hence further costs) for the recording of the patient's choice. On the other hand, in the unlikely event that such transfers cannot be legitimised, Watson for Oncology would not serve its support function despite the costs incurred by Dutch healthcare providers, which would in turn make data protection law an obstacle to healthcare provision.

Therefore, it cannot be stressed enough that efficient and uniform data protection rules in respect of health records are of paramount importance to cross-border healthcare and international cooperation in the health domain.

4.3. REGULATING CONTEXT – ACTORS, PURPOSES, PRINCIPLES AND NORMS

Controlling the purposes of processing, as well as the rules (norms), principles and obligations that the actors need to comply with, are key to preventing context transgression. Therefore, data protection legislation needs to regulate each of these elements of the context in order to efficiently prevent transgression consequences.

---

[148] There are exceptions under Article 49 GDPR for specific situations that require the transfer of a small number of health records, but transfers of big data sets of EHRs from a Member State to a UK-based processor, for example, would be much more difficult to justify.

[149] Schippers en Kamp tekenen brede Health Deal voor gerichte beslissingen in de kankerzorg, Rijksoverheid, 8 June 2016 < https://www.rijksoverheid.nl/actueel/nieuws/2016/06/08/schippers-en-kamp-tekenen-brede-health-deal-voor-gerichte-beslissingen-in-de-kankerzorg> Last accessed on 6 June 2018.

Turning to actors, unlike the DPD which only held data controllers (who determine the purposes and means of the processing)[150] liable, the GDPR also imposes specific obligations[151] on the processors[152] as well. Thus, for example, tech companies which apply advanced data analytics to health records will have to comply with specific sets of duties if they are controllers, joint controllers or just processors, respectively.[153] As it is a question of substance, and not of designation of roles, tech companies are likely to be held to the higher standard of liability of joint controllers when they dictate the technological means and/or specific purposes of the processing. This is intended to incentivise such companies to observe data protection principles and improve compliance with the new regulation compared to the old legal frameworks.

However, the Article 29 Working Party has previously acknowledged that there are inevitably many participating data controllers using the information in EHR systems, and so it recommended that a single institution be made responsible towards the data subjects for the proper handling of access requests.[154] Where that is the case, it may be up to the institution to use contractual terms and obligations to, in turn, hold tech companies accountable for compliance with the data protection principles. The drawback is that if a broad scope of processing purposes and activities is agreed (e.g. the initial DeepMind/NHS contract provisions), a public healthcare institution may end up liable for damages caused mainly by tech corporations acting as joint controllers.

One of the ways in which data protection legislation regulates the context of EHR analytics, which is discussed here, is by restricting the types of legitimate purposes for identifiable health data processing. This, in theory, should help prevent processing activities for inappropriate purposes, but in practice some exceptions are broad enough to allow it, as will be shown below.

Any information contained in an EHR is deemed to be sensitive personal information (Article 9 GDPR and Article 11 UKDPA) for the purposes of EU data protection law[155] and is therefore subject to a stricter regime of regulation. The protection appears to be strengthened under the

[150] Article 4(7) GDPR.
[151] Article 28 GDPR
[152] Article 4(8) GDPR.
[153] Articles 24-43 GDPR.
[154] WP29 131 at 7.
[155] WP29 131 at 7.

GDPR, which requires even certain stronger safeguards than the DPD to be set by Member States in national legislation on health data processing.[156]

There is a general prohibition of the processing of health data,[157] but exceptions are possible on several legal grounds. The first step of evaluating the efficiency and limits of data protection rules on processing of EHRs is examining the types of legal basis (general purposes) for the processing. Member States are free to impose further conditions and limitations in respect of the processing of genetic, biometric and health data,[158] so the Regulation's provisions must be read side by side with national law. In addition to supplementing the GDPR, the UKDPA makes minor amendments to various national legislation, including in respect of access to health records.[159]

4.3.1. The Role of Patient Consent, Determination and Control

The first potential legal ground for processing of EHRs is explicit consent,[160] which must be "freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her", as defined in Article 4(11) GDPR. Further conditions for consent have been introduced in Article 7 GDPR (there were no equivalent provisions under the old DPD regime). Consent can only be a valid legal ground for processing when the individual data subject has a genuine free choice (independent of social, financial, psychological *etc.* pressure) and is subsequently able to withdraw the consent[161] easily and without detriment.[162] When consent to the processing of data which is non-essential to the provision of a service is nonetheless made a condition for the provision of that service, the

[156] The legal framework and guidance on data protection under the Cross-border eHealth Information Services (CBeHIS) T6.2 JAseHN (draft v2 20/10/2016), eHealth Network < https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20161121_co18_en.pdf> Last accessed on 20 May 018.
[157] Article 9(1) GDPR.
[158] Article 9(4) GDPR.
[159] Access to Health Records Act 1990, Access to Medical Reports Act 1988, Access to Health Records (Northern Ireland) Order 1993 (S.I. 1993/1250 (N.I. 4)) *etc.*
[160] Article 9(2)(a) GDPR.
[161] Article 7(3) GDPR.
[162] WP29 131 at 8.

consent is not freely given.[163] In a medical context, the threat of non-treatment or lower quality treatment is considered coercion, which renders the consent invalid.[164] As a consequence, this legal ground cannot be used to legitimise processing by a health professional when the medical situation necessitates it.[165]

The Article 29 Working Party has already acknowledged that there may be practical difficulties in obtaining consent for the processing of patient records. However, Article 7(1) GDPR imposes a duty on the data controller relying on consent to be able to evidence it. If either one is impossible, then consent would be an invalid and potentially inappropriate legal basis.

Furthermore, consent must be specific.[166] This can be a particularly challenging requirement if the processing involves data mining algorithms and unsupervised deep learning techniques, which may rely on clustering and density estimation to find interesting properties in the datasets,[167] and may not have a pre-determined purpose. That is why machine learning is considered to be so opaque that it is debateable whether even the programmer who writes the code can predict how the software will arrive at a conclusion. Consent to processing without a specific purpose is unlikely to meet the requirement for specificity. Where the consent is given as a part of a written declaration concerning, *inter alia*, other matters, it must be distinguishable from the other matters.[168] For example, if a patient signs a written declaration of consent regarding a health-related service, there must be a clear distinction between consent to the service and the processing of the data.

But the most problematic requirement, when relying on this legal ground to legitimise processing with advanced computational techniques, is that the consent must be informed.[169] What this entails is an appreciation and an understanding of the facts and consequences of consenting. The patient must be given, in a clear and plain language, accurate and full

[163] Article 7(4) GDPR.
[164] WP29 131 at 8.
[165] WP29 131 at 8.
[166] Article 4(11) GDPR.
[167] Benjamin Shickel, Patrick J. Tighe, Azra Bihorac, and Parisa Rashidi, 'Deep EHR: A Survey Of Recent Advances In Deep Learning Techniques For Electronic Health Record (EHR) Analysis', (2017) IEEE Journal of Biomedical and Health Informatics PP(99), June 2017.
[168] Article 7(2) GDPR.
[169] Article 4(11) and Article 7(3) GDPR.

information of all relevant details, including the types of data processed, the purposes, recipients of possible transfers, his/her rights as the data subject, including the consequences of declining consent.[170] Providing adequate information about machine learning and data mining to lay persons, in such a way as to enable them to understand the process and appreciate the consequences, is not just a challenge, but may well be impossible[171] due to the "black box" (opaque) nature of these cutting-edge data analytics methods.

Data controllers processing information in EHR systems must provide certain information to data subjects, such as information on the identity of the controller, on the purposes of the processing, the types of data processed, on the recipients of the data and on the existence of a right of access and right of withdrawal. If the patient records are used for decisions based solely on automated processing, then information about such use would also have to be provided.[172] It is, however, highly unlikely, in the context of healthcare, for decisions to be based solely on data analytics without ultimate human intervention. Despite great advances in the field of healthcare informatics, AI and machine learning tools are still intended to assist doctors in medical diagnosis and help researchers with EHR analysis, but not to take away their decision-making autonomy. Indeed, AI may uncover fascinating correlations within big data sets, but it is up to data analysts and researchers to formulate hypotheses that explain them.

Having in mind all of the above requirements, consent appears to be an inappropriate legal ground[173] for most projects involving the advanced data analysis of EHRs, even though in many countries, *e.g.* Belgium[174] and the UK,[175] access to and processing of health records is still geared around consent. Furthermore, Recital 43 GDPR indicates that public authorities cannot rely on consent for processing when there is no genuine choice. Although the requirement for freely given, informed and specific consent *prima facie* enables the patient to agree only to

[170] WP29 131 at 9.
[171] Greenhalgh et al. "Patients' attitudes to the summary care record and HealthSpace: qualitative study." BMJ. 2008 Jun 7;336(7656):1290-5. doi: 10.1136/bmj.a114. Epub 2008 May 29. (2008), p. 1290.
[172] Article 22 (2) GDPR.
[173] See Sharon (N5), at 6.
[174] Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013) at 29.
[175] *Ibid* at 64; Julia Powles and Hal Hodson, "Google DeepMind and healthcare in an age of algorithms" Health Technol. (2017) < https://doi.org/10.1007/s12553-017-0179-1> Last accessed on 4 June 2018.

processing that he/she feels comfortable with (and thus avoid negative reactions on an individual level), there are practical considerations that cannot be ignored. For instance, there are patients who do not have legal capacity to provide valid consent, *e.g.* patients with Alzheimer's disease, and additional requirements will have to be met in respect of consent for processing of health records of minors.[176] Furthermore, EU or Member State law may not allow for the general prohibition in Article 9(1) GDPR to be lifted by the data subjects by virtue of consent. Therefore, in the context of healthcare provision, a different legal basis may be more suitable (from a legal perspective) to justify advanced data analytics.

But even when a different legal basis is used, self-determination should be of the paramount importance in the functioning of EHR systems.[177] Giving the patient autonomy for self-determination and control of the relevant EHR is intended to prevent negative perceptions and distrust. The Article 29 Working Party has stressed that the right to opt-out can also act as a different safeguard to the data subject's rights when explicit consent is not required.[178] Thus, the data subject would ultimately have a degree of control over the use of his/her health data.

Turning to exceptions to the requirement for explicit consent, one of the domestic derogations introduced in the UK within the medical information governance architecture—the so-called Caldicott principles and guidelines,[179] is the implied consent exception in the case of a direct care relationship. It became the object of much academic scrutiny and media attention in respect of the initial Google DeepMind/Royal Free agreement, because it sparked fears that the contracting parties wanted to circumvent the twin principles of transparency and patient self-determination for the sake of private interests. Although subsequently it was announced that there was a possibility for opt-out of the data sharing agreement, notifying patients after processing activities have been carried out is incompatible with data protection laws.[180] Indeed, it can be a cold comfort for data subjects to opt-out of processing if big data algorithms have

[176] Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013) at 65.
[177] WP29 131 at 13-14.
[178] *Ibid.*
[179] Julia Powles and Hal Hodson, "Google DeepMind and healthcare in an age of algorithms" Health Technol. (2017) at 9 < https://doi.org/10.1007/s12553-017-0179-1> Last accessed on 4 June 2018.
[180] See Articles 10 and 11 DPD, Articles 13 and 14 GDPR.

already mined their EHRs for similarities and patterns, and potentially made new medical discoveries on the basis of the data sets. The potential harms from such unauthorised medical analysis may not be sufficiently addressed through judicial remedies or financial compensation.

4.3.2. Vital Interests of the Data Subject or of another Natural Person

Article 9(2)(c) GDPR provides another legal ground which may be relied upon for the processing of health records which is "necessary to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent". The Article 29 Working Party has stressed the point that this exception cannot legitimise processing personal health data for purposes other than treatment of the data subject such as, for example, to carry out general medical research that will not yield results until some time in the future.[181] It has also recommended that this exception be applied only where the first consent of the two-step model has been given.[182] Limiting the purpose of the processing to direct patient care in life or death situations ensures that the values of the context remain the same, on the one hand, and significantly narrows the scope of application, on the other. It is traditionally envisaged that this ground could legitimise processing of health data only in exceptional cases, rather than justify computer-driven analysis of big data sets.

Although Recital 46 GDPR suggests that there may be circumstances, *e.g.* when the processing is necessary for the monitoring of epidemics or man-made disasters, in which either the 'vital interests' ground, or the 'public interest' ground, can legitimise the processing activities, this exception is unlikely to be relied on for big data analytics.

Nevertheless, the law firm Linklaters, which conducted a third party audit of the Google DeepMind/NHS agreement, has recently recommended using precisely the "vital interests" exception to legitimise the processing of patient records in the Streams app.[183] There are no supplementary provisions or restrictions in the UKDPA in respect of the vital interest exception in Article 9(2)(c) GDPR, but it remains to be seen if the UK Information Commissioner will

[181] WP29 131 at 9.
[182] Article 29 Working Party, Working Document 01/2012 on epSOS, WP 189 adopted on 25.01.2012, at 8.
[183] Natasha Lomas, 'Audit of NHS Trust's app project with DeepMind raises more questions than it answers', (*TechCrunch*, 13 June 2018) < https://techcrunch.com/2018/06/13/audit-of-nhs-trusts-app-project-with-deepmind-raises-more-questions-than-it-answers/?guccounter=1> Last accessed on 19 July 2018.

accept its use for the processing of "big data", or once again deem it to be an inappropriate legal basis for Streams.

4.3.3. Substantial Public Interest

Article 9(2)(g) GDPR (previously Article 8(4)DPD) postulates that when the processing is necessary for the purposes of substantial public interest, there may be further exceptions in Member State or Union law to the general prohibition for processing of sensitive data. The new provision clarifies that in addition to necessity, there are requirements of proportionality, respect for the essence of data protection and suitable safeguards for the fundamental rights of the data subjects. This provision is flexible enough to allow Member States to prioritise national policies, *e.g.* stimulating innovation in healthcare, while balancing them with the protection of the data subjects' rights through further legislation.

Section 10(3) UKDPA makes the reliance on the derogation legitimate only if the processing meets a condition in Part 2 of Schedule 1. There are twenty-four conditions, which provides a wide scope for processing. When the Data Protection Bill was first introduced, a privacy watchdog criticised the lack of a precise definition of "substantial public interest" or of an explanation why the seventeen conditions in the bill represented such interests.[184] Although these concerns were emphatically stated in a briefing to the House of Lords in Parliament and a letter to the UK Minister of State for Digital, the desired outcome was not achieved. Instead, further conditions were introduced, which thus expanded the scope of the derogation.

Nevertheless, the wording of certain conditions does seem to address specific context transgression issues. A peculiar condition is the one in Section 8 of Part 2 of Schedule 1 UKDPA, which allows processing of, *inter alia*, health data, for the purpose of promoting equality of opportunity or treatment. It expressly legitimises processing aimed at recognising bias towards groups of "people with different states of physical or mental health", presumably by identifying such groups for the sake of positively enforcing their rights. However, this may match one of the nightmare scenarios for many patients, especially when the processor is also

[184] Privacy International Makes Recommendations To Strengthen UK Data Protection Bill, 1 October 2017 <https://privacyinternational.org/press-release/626/privacy-international-makes-recommendations-strengthen-uk-data-protection-bill> Last accessed on 4 June 2018.

an employer like Amazon, which recently added the healthcare joint venture to the list of its many associated companies. Section 8(4) takes this into consideration, by prohibiting the processing if it is likely to cause distress or damage to an individual. Section 8(3) further prohibits processing which is intended to result in measures or decisions regarding a particular data subject. It can be argued that if the processing recognises bias towards patients with a specific condition, any action subsequently taken to promote equality should affect individual data subjects in that group. Section 8(5) implies respect for self-determination, despite the absence of consent, as it provides an opt-out mechanism. Like other conditions in Part 2 of Schedule 1 UKDPA, Section 8 provides safeguards against certain context transgression consequences, but not all.

4.3.4. Processing Data for Preventive or Occupational Medicine, Medical Diagnosis

One of the notable changes in the GDPR from the DPD is in respect of the legal basis for processing activities for preventive or occupational medicine.

Article 9(2)(h) GDPR[185] demands that the processing must be required and limits the legitimate purposes for the processing. The old Article 8(3) DPD legitimised processing which took place "for the purposes of preventive medicine, medical diagnosis, the provision of care or treatment or the management of health-care services", whereas Article 9(2)(h) GDPR adds to that list "the purposes of [...] occupational medicine, […] the assessment of the working capacity of the employee". Academics have previously noted that the legitimate purposes in the DPD provision were too broad,[186] so increasing the number of legitimate purposes in the GDPR is sure to attract new criticisms. But the new Article 9(2)(h) GDPR also uses more specific language than the old provision, justifying processing for "the provision of health or social care or treatment or the management of health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional".

[185] Like the old Article 8(3) DPD.
[186] Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013) at 68.

The final condition relating to the processors has also changed. Whereas under Article 8(3) DPD the data had to be "processed by a health professional subject […] to the obligation of professional secrecy or by another person also subject to an equivalent obligation of secrecy", Article 9(3) GDPR demands that the data be "processed by or under the responsibility of a professional subject to the obligation of professional secrecy under Union or Member State law or rules established by national competent bodies or by another person also subject to an obligation of secrecy under Union or Member State law or rules established by national competent bodies." The new provision seems to expand the scope of this legal basis and, compared to the DPD, is much more lenient in respect of the types of actors who may process the data, while holding the confidentiality-bound professional accountable for the processing.

Section 11(1) UKDPA reiterates the obligation of secrecy referred to in Article 9(3) GDPR. The Section restricts the scope of the first limb of the requirement by clarifying that the professional subject must be either a "health professional or a social work professional" (listed in Section 204 UKDPA), but seemingly expands the scope of the second limb: "[...] another person who in the circumstances owes a duty of confidentiality under an enactment or rule of law." Whether or not there is really a meaningful distinction between the scope of the GDPR article and the UKDPA Section may be determined in future case law of the respective courts.

Section 10(2) UKDPA states that a processing activity justified under point (b), (h), (i) or (j) of Article 9(2) of the GDPR can be deemed to have authorisation only if it "meets a condition in Part 1 of Schedule 1" UKDPA, namely if the processing is necessary for employment, social security and social protection, health or social care purposes, public health, research *etc*. As the UKDPA does not specify whether it has to meet a corresponding condition, it appears that there may be a mismatch between the GDPR exception and the UKDPA authorisation to use the exception, *e.g.* relying on the Article 9(2)(h) GDPR ('Preventive or occupational medicine') derogation on the basis of Section 1 of Part 1 of Schedule 1 UKDPA ('Employment, social security and social protection'). Despite that, the Article 29 Working Party held the view that processing operations in areas such as public health and social protection should be outside the

scope of application of Article 8 (3) DPD,[187] so it is likely that the EDPB will adopt the same position.

Interestingly, Part 1 of Schedule 1 UKDPA mostly repeats the language of Article 9 GDPR, but without introducing further conditions or providing specific definitions/examples of the listed types of legitimate purposes. This keeps the scope of the derogations quite broad. The broader the scope of the derogations, the easier it will be to justify further processing for secondary purposes, which may not be limited to the healthcare context and have serious implications for all stakeholders.

For example, machine-learning analytics of EHRs for the purpose of assessing the working capacity of employees may have a legal basis under Article 9(2)(h) GDPR or Section 2(2)(b) of Part 1 of Schedule 1 UKDPA if carried out under the responsibility of a confidentiality-bound professional. If the analysis reveals, *e.g.* that a group of employees have been treated for vitamin D deficiency, which is often viewed as an early predictor of multiple sclerosis[188] or depression,[189] then all members of that group may be deemed at risk to develop such conditions and receive unfavourable assessments of their working capacity. This may constitute prejudice towards every member of the group, but it is more difficult for members of groups with one or several shared genetic/physiological traits to establish such prejudice, than it is for members of self-aware groups explicitly protected by the law.

This specific scenario of employers discriminating on the basis of health data can be partially addressed through two provisions in the UKDPA. Section 184(1) UKDPA explicitly prohibits individuals from coercing a potential employee, an employee or a contractor into providing access to health records (subject to exceptions), and Section 185 UKDPA makes contractual clauses requesting such access void. Thus the UK legislation can address specific problems on the national level which the GDPR does not solve on the supranational level. It should, however,

---

[187] WP29 131 at 10.
[188] See, e.g. A. Ascherio, K.L. Munger, R. White, et al., 'Vitamin D as an Early Predictor of Multiple Sclerosis Activity and Progression' JAMA Neurol. 2014;71(3):306–314. doi:10.1001/jamaneurol.2013.5993; K. L. Munger, S. M. Zhang, E. O'Reilly, M. A. Hernán, M. J. Olek, W. C. Willett, A. Ascherio, "Vitamin D intake and incidence of multiple sclerosis", Neurology Jan 2004, 62 (1) 60-65; DOI: 10.1212/01.WNL.0000101723.79681.38
[189] D.J Armstrong, G.K Meenagh, I. Bickle, et al. Clin Rheumatol, "Vitamin D deficiency is associated with anxiety and depression in fibromyalgia" (2007) 26: 551. <https://doi.org/10.1007/s10067-006-0348-5> .

be noted that the GDPR is intended to set a standard of data protection and EU Member States are free to introduce certain derogations and further restrictions, including provisions similar to Sections 184 and 185 UKDPA.

4.3.5. Public Interest in the Area of Public Health

One of the new potential derogations from the general prohibition for processing of sensitive data is introduced in Article 9(2)(i) GDPR, which allows processing in the interests of public health. This article stipulates that there must be a legal basis in EU or Member State law, as well as specific safeguards for the rights and freedoms of data subjects. Examples of possible legitimate purposes are given in the article: "protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices". This will likely be the most widely used derogation in respect of legitimising secondary-purpose processing of EHRs by tech giants.

As noted above, Section 10(2) UKDPA refers to the conditions in Part 1 of Schedule 1 UKDPA, of which the public health or the health and social care conditions could potentially be met in the case of developing a clinical alert app like Streams. The definition of a medical device in EU law is fairly broad,[190] and it includes software used for specific medical purposes. It may therefore be possible to justify processing of health records for the purposes of improving healthcare AI developed by tech companies such as DeepMind, provided that there is authorisation in EU or Member State law.

The majority of patients would not reasonably expect their health records to be analysed by tech companies for the purposes of improving the quality of medicine or machine learning-algorithms. Many data subjects would object to their EHRs being shared with pharmaceutical companies or tech giants, so by association they may not agree to allow processing in the interests of such commercial entities. However, the public health derogation precludes data subjects from exercising self-determination through consent or opt-out mechanisms. On the

[190] For the currently applicable definition, see Art. 1(2)(a) of Council Directive 93/42/EEC on Medical Devices (MDD) (1993); for the definition after reform that will be applicable from 25 May 2020, see Art. 2(1) of Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC.

other hand, in the absence of self-determination, the informational burden can be equally distributed in public national healthcare systems. To truly improve public health and the healthcare services, however, there must also be a successful balance between the benefits for pharmaceutical companies or tech companies and their obligations to the patients and the healthcare providers.

4.3.6. Public Interest, Scientific or Historical Research or Statistical Purpose

Article 9(2)(j) GDPR provides another new exception to the prohibition of processing of sensitive data, in cases when the processing takes place for the purposes of public interest, scientific or historical research or statistics. In addition to the cumulative conditions of necessity, proportionality, respect for data protection and fundamental rights of the data subjects (all four of these conditions also apply to Article 9(2)(g) GDPR discussed above), further safeguards are introduced in Article 89, notably data minimization methods such as pseudonymisation and anonymisation (when possible). The efficiency of these privacy safeguards will be examined in further detail below.

By virtue of Section 10(2) UKDPA, as previously observed, processing relying on the Article 9(2)(j) GDPR derogation is deemed to have authorisation only if it "meets a condition in Part 1 of Schedule 1" UKDPA, namely if the processing is necessary for employment, social security and social protection, health or social care purposes, public health, research *etc*.

The wide scope of this derogation will make it easier for data controllers to justify data analytics of EHRs for a variety of purposes. It should also be emphasized that any type of processing may have a specific legitimate purpose, such as medical research, but that does not necessarily preclude it from serving multiple purposes, *e.g.* the development of data mining software that is subsequently made commercially available, or drug discovery through virtual clinical trials.[191] Even if the requirements to specify purposes for related processing activities in addition to the broad heading of 'research' are met,[192] the very method of conducting medical

[191] Lada Leyens et al., "Use of big data for drug development and for public and personal health and care", Official journal of the International Genetic Epidemiology Society, Volume 41, Issue1, January 2017, Pages 51-60, <https://doi.org/10.1002/gepi.22012> Last accessed on 8 June 2018.
[192] Article 29 Working Party, Opinion 03/2013 on purpose limitation WP 203 (02.04.2013) at 16.

research through machine learning can result in the development and improvement of software. While medical research, software and medicine development by private corporations certainly must be encouraged, regulation should address the need for the positive outcomes of data analytics to benefit the data subjects whose EHRs have been processed in these pursuits. Healthcare institutions and providers that join forces with data analytics companies should also be wary of the dangers of lock-in situations when services and software thus developed are leveraged against them for profit.

4.4. DATA PROTECTION PRINCIPLES

The data protection principles in Article 5 GDPR that general processing activities need to comply with can also have an indirect effect on context integrity.

For instance, the role of "lawfulness, fairness and transparency", which are the cornerstones of earning the public's trust in a healthcare system, is emphasized in Article 5(1)(a) GDPR. The old provision[193] only mentioned the first two requirements, but the addition emphasizes the strengthened role of transparency in data processing projects. Enforcement of this principle can therefore prevent scepticism, fear of stigmatisation and the potential 'chilling effect' they can have on individuals and groups of patients.

Further, if data are safely deleted from an EHR after the end of a set storage period, then undoubtedly the risks to the patient's privacy are mitigated. The storage limitation principle[194] allows personal data to be kept in a form that permits identification of data subjects for no longer than necessary for the processing purposes. On the other hand, it permits the storage of data after the expiry of the time limit, provided that they are used for "archiving purposes in the public interest, scientific or historical research purposes or statistical purposes in accordance with Article 89(1)". This is too broad a scope to efficiently restrict the retention period.

The newly introduced obligation of the data controller to ensure the integrity and confidentiality of the data[195] is an especially valuable guiding principle in the context of health data processing.

[193] Article 6(1)(a) DPD.
[194] Article 5(1)(e) GDPR.
[195] Article 5(1)(f) GDPR.

The Article 29 Working Party has nevertheless recognised that the confidentiality of health records is not entirely within the control of medical professionals and that has a negative impact on patients' trust in the records system.[196]

4.4.1. Purpose Limitation

The principle of purpose limitation[197] is one of the chief data protection principles which applies to personal data in EHRs.[198] As defined by the Article 29 Working Party in its opinion on the purpose limitation principle, it is designed to "prevent the use of individuals' personal data in a way (or for further purposes) that they might find unexpected, inappropriate or otherwise objectionable,"[199] while allowing processing for uses which are not incompatible (within the same context). This principle is clearly intended to prevent context transgression, closely resembling Nissenbaum's ideas, and it is notable that part of the Opinion focused on "Context and Strategic Consequences".[200] The Working Party has also emphasized that purpose limitation is linked to the principles of transparency, predictability and user control.[201] All of these are vital safeguards which can preserve context integrity.

Nevertheless, while the initial collection of data in EHRs can easily satisfy the requirements for specified, explicit and legitimate purposes, further processing for secondary purposes can be more controversial. The Article 29 Working Party's view was that if the secondary purpose of the processing is "incompatible with the purposes specified at collection is unlawful and therefore not permitted."[202] As it is up to the data controllers to make the assessment of incompatibility of purposes, it is largely up to the respective data protection authorities to ensure enforcement of the principle in practice, and the respective courts to address disputes and provide remedies.

However, the GDPR seems to undermine the application of the safeguards of purpose limitation and storage limitation. Article 5(1)(b) GDPR explicitly declares "further processing for

[196] WP29 131 at 20.
[197] It was partially embodied in Article 6(1)(b) of the DPD, now replaced by Article 5(1)(b) GDPR.
[198] WP29 131 at 6.
[199] Article 29 Working Party, Opinion 03/2013 on purpose limitation WP 203 (02.04.2013) at 11.
[200] WP29 203 at 14.
[201] WP29 203 at 13-14.
[202] WP29 203 at 36.

archiving purposes in the public interest, scientific or historical research purposes or statistical purposes [...] in accordance with Article 89(1)" not incompatible with the purpose for collection. Critics have attacked this new research exemption in the GDPR as a step too far and have warned of the dangers specifically in respect of health data and especially genetic data.[203] Furthermore, in light of the exception from the storage limitation principle,[204] the research derogation allows almost free reign over data for unspecified periods of time.

4.4.2. Data minimization

Article 5(1)(c) GDPR places an obligation on the data controller to minimise data collection to an adequate level regarding the purposes of processing. Anonymisation or pseudonymisation are helpful methods for mitigating the risks to the data subjects' privacy, but they are not perfect solutions as they do not address the risks to group privacy.

4.4.2.1. Anonymisation

According to its Recital 26, the GDPR does not apply to information which does not relate to an identified or identifiable natural person or to personal data rendered modified in such a way that the data subject is not or no longer identifiable. The EU does not provide for a standard of successful anonymisation,[205] but there has been some guidance on the matter by the Article 29 Working Party.[206]

However, there is always a risk factor inherent to anonymisation,[207] and there is much academic discourse on the threat of re-identification, *e.g.* through linkage attacks[208] (as mentioned in the

---

[203] See, *inter alia*, Kärt Pormeister; Genetic data and the research exemption: is the GDPR going too far?, International Data Privacy Law, Volume 7, Issue 2, 1 May 2017, Pages 137–146, https://doi.org/10.1093/idpl/ipx006; B. Custers and H. Uršič, 'Big data and data reuse: a taxonomy of data reuse for balancing big data benefits and personal data protection', International Data Privacy Law, 6 (1) (2016) 4-15; Menno Mostert et al. "From Privacy to Data Protection in the EU: Implications for Big Data Health Research", European Journal of Health Law, Volume 25, Issue 1, 2017, 43 – 55.
[204] Article 5(1)(e) GDPR.
[205] Axel von dem Bussche; Paul Voigt, *The EU general data protection regulation (GDPR): a practical guide*, (Springer International Publishing 2017) at 14.
[206] See Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques (European Commission Working Paper No. 216, 0829/14/EN, 2014).
[207] Axel von dem Bussche; Paul Voigt, *The EU general data protection regulation (GDPR): a practical guide*, (Springer International Publishing 2017) at 14.
[208] Further discussed in Solon Barocas, Helen Nissenbaum, "Big Data's End Run around Anonymity and Consent" pp.44-75 in Julia Lane et.al. "Privacy, Big Data, and the Public Good: Frameworks for Engagement",

previous chapter). The GDPR, as did the DPD before it, limits the identifiable concept by a reasonableness standard in Recital 26, *i.e.* to the extent that only all the means likely reasonably to be used to identify someone are taken into account.

To prevent harms and mitigate the threats from the processing of anonymised data, the UKDPA goes further than the GDPR. As an additional safeguard, it criminalises attempts to re-identify 'de-identified' information,[209] which is defined as information "processed in such a manner that it can no longer be attributed, without more, to a specific data subject" (such as through encryption, anonymisation or pseudonymisation), but there are several defences and exceptions.[210]

Moreover, while the GDPR does not apply to data after anonymisation and does not seem to address the consequences of further processing, the UKDPA introduces new safeguards in Section 170 UKDPA against offences such as unlawful obtaining of personal data or unlawfully retaining them "without the consent of the person who was the controller in relation to the personal data when it was obtained".[211] It is a further offence if the data so obtained or retained are sold on to a third party. Although it is not clear whether companies and other private commercial entities can be held similarly liable, or only natural persons, in theory it should prevent unauthorised attempts to profit from the processing of EHR datasets and subsequent harms.

But even when health data are successfully anonymised and no re-identification is attempted, the data sets of EHRs are extremely valuable for a variety of purposes, including medical research, data analytics,[212] healthcare AI development (such as DeepMind's Streams app),

---

(Cambridge University Press 2014). See also: Jane Henriksen-Bulmer and Sheridan Jeary (Dr), "Re-identification attacks—A systematic literature review", International Journal of Information Management, Volume 36, Issue 6, Part B, December 2016, Pages 1184-1192, < https://doi.org/10.1016/j.ijinfomgt.2016.08.002> Last accessed on 8 June 2018; Shouling Ji, Prateek Mittal, Raheem Beyah, "Graph Data Anonymization, De-Anonymization Attacks, and De-Anonymizability Quantification: A Survey", IEEE Communications Surveys & Tutorials (Volume: 19, Issue: 2, Second quarter 2017), 1305 – 1326.

[209] Section 171 UKDPA.

[210] Sections 171 and 172 UKDPA.

[211] Section 170(1)(c) UKDPA.

[212] Adeel Anjum et al., "An efficient privacy mechanism for electronic health records", Computers & Security Volume 72, January 2018, Pages 196-211, <https://doi.org/10.1016/j.cose.2017.09.014> Last accessed on 8 June 2018.

statistics, healthcare system management and policy, drug discovery and development[213] *etc.* This suggests that big data analytics can indeed have many benefits when restricted to the same specific healthcare context, and supports the argument that strict regulation of EHR analytics does not unduly stifle innovation.

4.4.2.2. Pseudonymisation

The new concept of 'pseudonymisation' is introduced in Article 4 GDPR. It is defined as 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. Unlike anonymisation, personal data that have undergone pseudonymisation are explicitly defined to remain personal data under EU data protection laws.[214] Thus, pseudonymised data are subject to the safeguards in the GDPR, as well as domestic legislation such as the UKDPA in the UK.

The new offence in Section 171 UKDPA of re-identifying 'de-identified' information applies to pseudonymised data as well, adding a useful safeguard for the prevention of context transgression harms.

A novel approach to sensitive data management through pseudonymisation, called Polymorphic Encryption and Pseudonymisation (PEP),[215] can enhance the legal protection through technical solutions against context transgression. The method, which is still in development, restricts access to the encrypted and pseudonymised EHR (or parts of it) only to specifically authorised users for specifically authorised purposes, thereby preventing unauthorised/inappropriate data flows (including unauthorised access by the storage facility).[216] In future, it would be preferable for PEP to be standardised (as a pseudonymisation method for EHRs), as well as recommended

---

[213] Lada Leyens et al., "Use of big data for drug development and for public and personal health and care", Official journal of the International Genetic Epidemiology Society, Volume41, Issue1, January 2017, Pages 51-60, <https://doi.org/10.1002/gepi.22012> Last accessed on 8 June 2018.
[214] Article 4(5) GDPR.
[215] See, *e.g.* Eric Verheul, Bart Jacobs, Carlo Meijer, Mireille Hildebrandt, Joeri de Ruiter, "Polymorphic Encryption and Pseudonymisation for Personalised Healthcare: A Whitepaper, Version 1.1" (Institute for Computing and Information Sciences, Radboud University Nijmegen, The Netherlands).
[216] *Ibid* at 8.

by legislation, as it is the combination of the different modalities of regulation (law and architecture, in addition to standards) that can offer the stronger protection than each of them in isolation.

4.5. ADDRESSING CONTEXT TRANSGRESSION HARMS

Non-compliance with the GDPR is dealt with in remedies, liability and penalties provisions,[217] which are similar to those in the DPD. Nevertheless, under the GDPR there are specific rights of data subjects with a wider scope than those in the old regime. Furthermore, the fines for non-compliance which can be imposed under Article 83 GDPR can be up to 20,000,000 EUR or 4% of total worldwide turnover. Thus processors and data controllers alike shall be motivated to comply under threat of financial penalties.

In addition to compensation to the data subjects[218] for material or non-material damage (context transgression harms), there is also a right to an effective judicial remedy against a controller or processor.[219] Whereas the DPD enforced liability only on the data controller, under the GDPR, compensation can be claimed from both the processor and the controller. The difficulty with claiming compensation for advanced data analytics of EHR datasets will clearly be in establishing and quantifying harm to an individual, but even more so to groups which may not be self-aware and legally defined.

Article 84 GDPR allows Member States to make provisions for further penalties. The UKDPA, accordingly, deals with enforcement in Part 6 UKDPA. As highlighted above, it creates new criminal offences in respect of personal data, such as re-identification of personal data without consent of the controller, and it also modernises criminal offences previously contained in the old Data Protection Act 1998, such as the offence of unlawful obtaining of personal data which is extended to include unlawful "retention" of data without the data controller's consent (even if the data was initially obtained lawfully). These safeguards undoubtedly strengthen the protection of context integrity on the national level.

[217] Articles 77-84 GDPR.
[218] Article 82 GDPR.
[219] Article 79 GDPR.

It follows that data protection regulation in the EU and the UK primarily aims to prevent context transgression, rather than address the consequences from it. Many harms remain outside the scope of data protection law, as illustrated by the examples in the previous chapter, and necessitate the use of other regulatory instruments to address the damage.

CHAPTER 5: CONCLUSION

All of the risks highlighted in the previous chapters should not imply that there could not be multiple benefits for individuals, efficiency gains, medical innovations and overall improvements of the quality of healthcare as a result of EHR advanced analytics by tech companies. But the difficult question arises whether national and international regulations are sufficiently capable of safeguarding the rights and enforcing the obligations of the various stakeholders while enabling the successful implementation of innovative machine learning projects. Close scrutiny of their efficiency and a critical analysis of any gaps in them can be useful to provide guidance on their potential application and recommendations on how to improve them.

There are six main points that became apparent in the foregoing critical assessment of the GDPR and the UKDPA from a context transgression perspective.

Firstly, there is little that either the GDPR or the UKDPA can do to prevent loss of public trust and negative personal reactions when "big data" analytics are applied by ICT companies to patient records. Both introduce new derogations from the general prohibition to EHR processing (the public health and the research derogations in the GDPR, as well as any additional UK-specific amendments introduced by the Secretary of State), which may take data subjects by surprise. In contrast, being informed and asked for permission before processing takes place can help adjust patients' reasonable expectations somewhat, so as to avoid their dismay. But the specificity required for informed and explicit consent to legitimise processing, as well as the additional administrative and financial burden to obtain it, may make it unattractive for companies, unless it is the only legal ground allowing cross-border data transfers in the circumstances. The transparency principle in the GDPR can also help inspire public trust, however the potential outcomes of the processing may still be unpredictable and upset the patients, especially if their consent has not been obtained. Although the UKDPA attempts to prohibit processing which is likely to cause distress to the data subjects, this may be a difficult provision to apply in practice as it is open to different interpretations by the UK courts and Information Commissioner, and is also a matter of degree. One way to address the shortcomings of the GDPR and the UKDPA in this respect would be through information campaigns about AI in healthcare, so as to initiate a cultural change.

Secondly, the rights of groups, especially groups which are not self-aware and recognised in law (*e.g.* genetic groups, or those emerging in the course of data mining), cannot be efficiently protected by the GDPR or the UKDPA. Risks to group privacy are not the focal point of either instrument, leaving some scope for marginalisation of vulnerable groups. For instance, the greatest negative impact of context transgression may be felt by those who would require more expensive treatments, or those at a higher risk of developing a certain condition, when the purposes of processing are not compatible with the context of direct care. The GDPR aims to address this through its purpose limitation principle and the overarching principle of protection of the fundamental rights of natural persons. On the other hand, it also introduces broad derogations in the public interest (public health and research), which may contradict group privacy rights in some situation. The GDPR thus requires balancing of stakeholders' rights in each situation, which can subsequently be done by the CJEU. Until then, it will fall to the EDPB to provide guidance (updating the Article 29 Working Party's opinion in respect of EHR processing should be prioritised due to its significance for the EU's eHealth Action Plan[220]). The UKDPA, on the other hand, encourages processing aimed at the positive enforcement and protection of groups of people of different mental and physical states. This, however, necessitates close scrutiny of the precise purposes of the processing by the Information Commissioner of the UK.

Thirdly, the GDPR, as well as the UKDPA, have created legal uncertainty for all parties to which they respectively apply. The GDPR introduced higher liability (including considerable fines) for non-compliance than the DPD, and imposed new obligations for data controllers, joint controllers and even processors. Although these parties need to conduct self-assessments, the respective Data Protection Authority (or the EDPB) may disagree with their conclusions in the course of investigation of complaints. Thus a company claiming to be merely a processor can be deemed to be a (joint) data controller instead, therefore subject to additional obligations and liability. This increases the financial risks for companies, which in turn can pressure them to look for alternative revenue from re-purposing of EHRs (even in breach of the medical care

[220] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century' /COM/2012/0736 final <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012DC0736> Last accessed on 4 December 2017.

context). Nevertheless, if the EDPB instead adopts its predecessor's opinion that healthcare institutions should be held liable as the data controller (so as to facilitate handling of patient complaints and reduce the onus on the processing ICT companies), this could result in a disproportionate burden on (public) healthcare institutions for damages in cases when private tech corporations with all the expertise control the data and cause the harm. Again, guidance from the EDPB and the jurisprudence of the CJEU can address this in time, but the UK's complicated relationship with the EU may continue to bring up problems for its parallel enforcement of data protection.

Fourthly, the derogations introduced in the GDPR (especially the public interest/research exception) and implemented in the UKDPA (with possibility for further derogations approved by the British Secretary of State) are far too broad and likely to allow context transgression. The additional requirement for data pseudonymisation or anonymisation cannot eliminate the risk entirely. It is just the opposite when it comes to the latter, as there is no binding standard in respect of EHRs, and anonymised data are not covered by the GDPR or the UKDPA (so inappropriate flows of information outside the direct care context may be possible). The UKDPA's special new offences (*e.g.* de-anonymising data or transferring personal records without the data controller's authorisation) go further to prevent context transgression. However, pseudonymisation seems to be the best way to retain legal and combine it with technical protection, and the GDPR has already been praised for distinguishing it from anonymisation.[221]

Fifthly, the GDPR, and to a lesser extent the UKDPA, does not take full account of the nature of advanced data processing technologies such as machine learning and data mining. For instance, certain AI algorithms first need to be "fed" personal data (*e.g.* X-rays of patients) in order to be applied to big EHR data sets later and recognise similar conditions. Any software errors[222] and implicit biases in the algorithms can be passed on to research or medical devices *etc.* developed on the basis of the interim processing, but if the initial data is deemed to be

[221] See e.g. Stalla-Bourdillon, Sophie and Knight, Alison, "Anonymous Data v. Personal Data — A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data" (March 6, 2017). Wisconsin International Law Journal, 2017. < https://ssrn.com/abstract=2927945> Last accessed on 16 August 2018.
[222] "Bugs" or "glitches" that can cause drastic problems for the processing or produce false results.

anonymised before or after the processing the consequences will not be covered by data protection remedies. Due to UK's experience with EHR processing projects, the UKDPA does however go further than the GDPR and prohibits de-anonymising personal data to prevent "linkage attacks" on anonymised data.

To sum up, the GDPR, as a supranational regulatory instrument, can harmonise regulation throughout the EU, but fails to achieve its full potential for efficient regulation of multinational tech titans due to the broad new derogations it introduces (which seem to allow a degree of context transgression for the sake of stimulating innovation). This allows Member States to decide on the national level whether to strengthen or weaken the set standard of protection by using the derogations. The primary advantages of the EU regime lie in its uniform interpretation of the GDPR, its dispute resolution and enforcement mechanisms, as well as ensuring cross-border data flows.

On the other hand, the UKDPA, which is intended to supplement the GDPR on the national level, does present solutions to a few specific problematic scenarios (*e.g.* employers coercing employees to provide health data), which can better reflect the UK's considerable experience with AI projects in healthcare and societal attitudes and reasonable expectations. Nevertheless, it also has certain structural problems, resulting in *prima facie* loopholes in the protection against context transgression.

As demonstrated above, both the GDPR and the UKDPA have their limited scope of efficiency in addressing context transgression in the case of processing of EHRs by tech companies, so both require certain changes to strengthen the protection against it. As the GDPR is intentionally designed to be flexible and serve as a baseline to be supplemented soon (by Member States' legislation, EDPB's opinions and CJEU's guidance and jurisprudence), there are currently more unanswered questions as to the UK regime's future enforcement and amendments, as well as to the position of the healthcare stakeholders under it.

Bibliography

Primary Sources:

Council Directive 93/42/EEC on Medical Devices (MDD) (1993)

Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108) and the Additional protocol to Convention 108 regarding supervisory authorities and transborder data flows (ETS No. 181).

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)

Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC.

Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications).

United Kingdom, Data Protection Act 2018 c.12

CJEU cases:

Google Spain SL and Google Inc. v Agencia Española de Protección de Datos (AEPD) and Mario Costeja González (C-131/12).

Rechnungshof (C-465/00) v Österreichischer Rundfunk and Others.

Secondary Sources:

Books and contributions to books:

Axel von dem Bussche and Paul Voigt, *The EU general data protection regulation (GDPR): a practical guide*, (Springer International Publishing 2017) at 14.

Carlisle George, Diane Whitehouse and Penny Duquenoy, *eHealth: Legal, Ethical and Governance Challenges* (Springer 2013).

Helen Nissenbaum, *Privacy In Context:Technology, Policy, and the Integrity of Social Life*, (Stanford University Press, 2010).

Ian P. McLoughlin, Karin Garrety, and Rob Wilson, *The Digitalization of Healthcare Electronic Records and the Disruption of Moral Orders* (Oxford University Press, 2017)

Linnet Taylor, Luciano Floridi, Bart van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies*, (Philosophical Studies Series, Springer, 2017).

Nadezhda Purtova, 'Health Data for Common Good: Defining the Boundaries and Social Dilemmas of Data Commons' (9 July 2016). in Ronald Leenes, Nadezhda Purtova, Samantha Adams (eds.) (2017) *Under Observation - The Interplay Between eHealth and Surveillance*, (Springer ; Tilburg Law School Research Paper No. 15/2016). <https://ssrn.com/abstract=2807455> Last accessed on 10 June 2018.

Pradeep K. Sinha et al., *Electronic Health Record: Standards, Coding Systems, Frameworks, and Infrastructures* (1st edn. Wiley 2013).

Solon Barocas and Helen Nissenbaum, 'Big Data's End Run around Anonymity and Consent' pp.44-75 in Julia Lane et.al., *Privacy, Big Data, and the Public Good: Frameworks for Engagement*, (Cambridge University Press 2014).

Journal Articles:

A. Docherty, 'Big Data – Ethical Perspectives', (2014) Anaesthesia, 69, 387–398, 390 <https://onlinelibrary.wiley.com/doi/pdf/10.1111/anae.12656 > Last accessed on 20 March 2018.

Adeel Anjum et al., "An efficient privacy mechanism for electronic health records", Computers & Security, Volume 72, January 2018, Pages 196-211, <https://doi.org/10.1016/j.cose.2017.09.014> Last accessed on 8 June 2018.

Amir Gandomi, Murtaza Haider, 'Beyond the hype: Big data concepts, methods, and analytics', International Journal of Information Management, (2015), Volume 35, 137–144.

Andrew D. Murray, 'Data transfers between the EU and UK post Brexit?', (2017) International Data Privacy Law, Volume 7, Issue 3, 1 August 2017, Pages 149–164, https://doi.org/10.1093/idpl/ipx015 Last accessed on 20 May 2018.

Audrey Guinchard, 'Contextual Integrity and EU Data Protection Law: Towards a More Informed and Transparent Analysis' (6 March 2017). <https://ssrn.com/abstract=2946772> Last accessed on 10 May 2018.

B Chen, AJ Butte, 'Leveraging big data to transform target selection and drug discovery', (2016) Clinical Pharmacology & Therapeutics, 99 3, 285-297.

B. Custers and H. Uršič, 'Big data and data reuse: a taxonomy of data reuse for balancing big data benefits and personal data protection', International Data Privacy Law, 6 (1) (2016) 4-15.

Benjamin Shickel, Patrick J. Tighe, Azra Bihorac, and Parisa Rashidi, 'Deep EHR: A Survey Of Recent Advances In Deep Learning Techniques For Electronic Health Record (EHR) Analysis', (2017) IEEE Journal of Biomedical and Health Informatics PP(99), June 2017.

Craig Konnoth, 'Health Information Equity', 165 U. Pa. L. Rev. 1317 (2017), 1343 <http://scholar.law.colorado.edu/articles/701> Last accessed 2 February 2018.

Diane Whitehouse, Carlisle George, and Penny Duquenoy, 'eHealth: legal, ethical and governance challenges - an overview', Global Telemedicine and eHealth Updates: Knowledge Resources, Vol. 4, (2011, International Society for Telemedicine & eHealth (ISfTeH)), 423-428.

Eleni Entzeridoua, Evgenia Markopouloua, Vasiliki Mollaki, 'Public and physician's expectations and ethical concerns about electronic health record: Benefits outweigh risks except for information security', (2018) International Journal of Medical Informatics, Volume 110, February 2018, 98.

Fischbacher et al, 'Record linked retrospective cohort study of 4.6 million people exploring ethnic variations in disease: myocardial infarction in South Asians', BMC Public Health 2007 < https://doi.org/10.1186/1471-2458-7-142> Last accessed on 20 March 2018.

Gianclaudio Malgieri, 'Ownership' of Customer (Big) Data in the European Union: Quasi-Property as Comparative Solution? (November 20, 2016). Journal of Internet Law, Vol. 20, n.5, November 2016. <https://ssrn.com/abstract=2916079>.

Greenhalgh et al., 'Patients' attitudes to the summary care record and HealthSpace: qualitative study.' BMJ. 2008 Jun 7;336(7656):1290-5. doi: 10.1136/bmj.a114. Epub 2008 May 29. (2008), p. 1290.

Helen Nissenbaum, 'A Contextual Approach to Privacy Online' (2011). Daedalus 140 (4), Fall 2011: 32-48, 45 <https://ssrn.com/abstract=2567042> Last accessed on 25 March 2018.

James Rachels, 'Why Privacy Is Important.' Philosophy & Public Affairs Vol. 4, No. 4 (Summer, 1975), pp. 323-333.

Jane Henriksen-Bulmer and Sheridan Jeary (Dr), "Re-identification attacks—A systematic literature review", International Journal of Information Management, Volume 36, Issue 6, Part B, December 2016, Pages 1184-1192, < https://doi.org/10.1016/j.ijinfomgt.2016.08.002> Last accessed on 8 June 2018.

Jason Chung, 'What Should We Do About Artificial Intelligence in Health Care?' (30 January 2018). NYSBA Health Law Journal, Winter 2017, Vol. 22, No. 3. <https://ssrn.com/abstract=3113655> Last accessed on 18 March 2018.

Jason Chung, and Amanda Zink,'Hey Watson, Can I Sue You for Malpractice? Examining the Liability of Artificial Intelligence in Medicine' (23 November 2017). Forthcoming, Asia-Pacific Journal of Health Law, Policy and Ethics. <https://ssrn.com/abstract=3076576> Last accessed on 30 March 2018.

Julia Powles and Hal Hodson, "Google DeepMind and healthcare in an age of algorithms" (2017) Health and Technology, December 2017, Volume 7, Issue 4, 351–367 <https://doi.org/10.1007/s12553-017-0179-1> Last accessed on 4 December 2017.

K. Irion and G. Luchetta, 'Online personal data processing and EU data protection reform' (8 April 2013) CEPS Task Force. Report of the CEPS Digital Forum. Centre for European Policy Studies 2013, at 57, <http://ssrn.com/abstract=2275267> Last accessed on 11 May 2018.

Kärt Pormeister, 'Genetic data and the research exemption: is the GDPR going too far?', (2017) International Data Privacy Law, Volume 7, Issue 2, 1 May 2017, Pages 137–146, <https://doi.org/10.1093/idpl/ipx006> Last accessed on 1 June 2018.

Kenneth M Boyd, 'Ethnicity and the ethics of data linkage', BMC Public Health 2007 7:318 <https://doi.org/10.1186/1471-2458-7-318> Last accessed on 20 March 2018.

Khaled El Emam and Sam Rodgers, 'Anonymising and sharing individual patient data', (2015) The BMJ, <https://www.bmj.com/content/350/bmj.h1139.abstract> Last accessed on 1 June 2018

Kristin Lacy-Jones, Philip Hayward, Steve Andrews, Ian Gledhill, Mark McAllister, Bertil Abrahamsson, Amin Rostami-Hodjegana, Xavier Pepine, 'Biopharmaceutics data management system for anonymised data sharing and curation: First application with orbito IMI project', (2017) Computer Methods and Programs in Biomedicine, Volume 140, March 2017, Pages 29-44.

L. Dauwerse, T. A. Abma, B. Molewijk, G. Widdershoven, 'Goals of Clinical Ethics Support: Perceptions of Dutch Healthcare Institutions', (2013) Health Care Analysis, December 2013, Volume 21, Issue 4, pp 323–337, <https://link.springer.com/article/10.1007/s10728-011-0189-5> Last accessed on 30 March 2018.

Lada Leyens et al., "Use of big data for drug development and for public and personal health and care", Official journal of the International Genetic Epidemiology Society, Volume 41, Issue1, January 2017, Pages 51-60, <https://doi.org/10.1002/gepi.22012> Last accessed on 8 June 2018.

Menno Mostert et al., 'From Privacy to Data Protection in the EU: Implications for Big Data Health Research', European Journal of Health Law, Volume 25, Issue 1, 2017, 43 – 55.

N Szlezák, M Evers, J Wang, L Pérez, 'The Role of Big Data and Advanced Analytics in Drug Discovery, Development, and Commercialization', (2014) Clinical Pharmacology & Therapeutics; 95 5, 492–495.

Nadezhda Purtova, 'Illusion of Personal Data as No One's Property' (October 29, 2013), Law, Innovation, and Technology, Volume 7, Issue 1, 2015. < https://ssrn.com/abstract=2346693> Last accessed on 1 June 2018.

Neil Mehta, Murthy V.Devarakonda, 'Machine Learning, Natural Language Programming, and Electronic Health Records: the next step in the Artificial Intelligence Journey?', (5 March 2018), Journal of Allergy and Clinical Immunology, 2, <https://doi.org/10.1016/j.jaci.2018.02.025> Last accessed 25 March 2018.

Patricia Balthazar et al., 'Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics', (2018) Journal of the American College of Radiology, (March 2018 Volume 15, Issue 3, Part B, Pages 580–586) 584, < https://doi.org/10.1016/j.jacr.2017.11.035> Last accessed on 28 March 2018.

Patricia Balthazar et al., 'Protecting Your Patients' Interests in the Era of Big Data, Artificial Intelligence, and Predictive Analytics', (2018) Journal of the American College of Radiology, March 2018, Volume 15, Issue 3, Part B, Pages 580–586 <https://doi.org/10.1016/j.jacr.2017.11.035> Last accessed on 30 March 2018.

Petersen C, Adams SA, DeMuro PR., 'mHealth: Don't Forget All the Stakeholders in the Business Case.', (2015) Eysenbach G, ed. Medicine 20 2015; 4(2):e4. < https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4713907> Last accessed on 4 December 2017.

Roehrich, Jens and Lewis, Michael and George, Gerard, 'Are Public-Private Partnerships a Healthy Option? A Systematic Literature Review' (2014) Social Science & Medicine, Vol. 113, pp. 110-119. <https://ssrn.com/abstract=2955093> Last accessed on 1 June 2018.

Sally Wyatt, Anna Harris, Samantha Adams, Susan E Kelly, 'Illness Online: Self-reported Data and Questions of Trust in Medical and Social Research', (2013) Theory, Culture & Society, vol. 30, Issue 4, 131 – 150, 1.

Shouling Ji, Prateek Mittal, Raheem Beyah, "Graph Data Anonymization, De-Anonymization Attacks, and De-Anonymizability Quantification: A Survey", IEEE Communications Surveys & Tutorials (Volume: 19, Issue: 2, Second quarter 2017), 1305 – 1326.

Stalla-Bourdillon, Sophie and Knight, Alison, "Anonymous Data v. Personal Data — A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data" (March 6, 2017). Wisconsin International Law Journal, 2017. < https://ssrn.com/abstract=2927945> Last accessed on 16 August 2018.

Tamar Sharon, 'The Googlization of health research: from disruptive innovation to disruptive ethics', (2016) Personalized Medicine, Volume 13, Issue 613 Oct 2016.

Tjeerd-Pieter van Staa, Ben Goldacre, Iain Buchan, Liam Smeeth, 'Big health data: the need to earn public trust', (14 July 2016), BMJ : British Medical Journal; London Vol. 354.

Websites and online newspaper articles:

'Eleven of 14 NHS health boards hit by ransomware cyber-attack' (*BBC News*, 12 May 2017), <http://www.bbc.com/news/uk-scotland-39896639> Last accessed on 1 April 2018.

'Facebook under fire in escalating data row' (*BBC News*, 19 March 2018) <http://www.bbc.com/news/technology-43461865> Last accessed on 29 March 2018.

'Google DeepMind NHS app test broke UK privacy law', *BBC news*, <http://www.bbc.co.uk/news/technology-40483202> Last accessed on 12 October 2017.

'Google's Deepmind NHS deal 'inexcusable', says academic paper', *The Register* (16 March 2017), <https://www.theregister.co.uk/2017/03/16/googles_deepmind_and_royal_free_hospital_deal_inexcusable/> Last accessed on 12 October 2017.

A. Ascherio, K.L. Munger, R. White, et al., 'Vitamin D as an Early Predictor of Multiple Sclerosis Activity and Progression' JAMA Neurol. 2014;71(3):306–314. <doi:10.1001/jamaneurol.2013.5993>; K. L. Munger, S. M. Zhang, E. O'Reilly, M. A. Hernán, M. J. Olek, W. C. Willett, A. Ascherio, "Vitamin D intake and incidence of multiple sclerosis", Neurology Jan 2004, 62 (1) 60-65; DOI: 10.1212/01.WNL.0000101723.79681.38

Adam Kucharski, 'Google's flu fail shows the problem with big data', (The Conversation, 24 October 2013) <https://theconversation.com/googles-flu-fail-shows-the-problem-with-big-data-19363> Last accessed on 4 December 2017.

Annie Palmer, 'Amazon's secret health lab revealed: 'Grand Challenge' working on everything from curing cancer to using AI to analyse medical records', Daily Mail (5 June 2018) <http://www.dailymail.co.uk/sciencetech/article-5810087/Amazons-secret-health-lab-revealed-Grand-Challenge-working-cancer-research-medical-records.html> Last accessed on 1 July 2018.

Bruce Japsen, 'It's Official: Amazon Enters Pharmacy Business With PillPack Acquisition' , (*Forbes*, 28 June 2018), < https://www.forbes.com/sites/brucejapsen/2018/06/28/its-official-amazon-enters-pharmacy-business-with-pillpack-deal/#6035c07113fa> Last accessed on 1 July 2018.

Bruce Japsen, 'It's Official: Amazon Enters Pharmacy Business With PillPack Acquisition', (*Forbes*, 28 June 2018), < https://www.forbes.com/sites/brucejapsen/2018/06/28/its-official-amazon-enters-pharmacy-business-with-pillpack-deal/#6035c07113fa> Last accessed on 1 July 2018.

Clifton Leaf, 'Amazon–JPMorgan–Berkshire Hathaway: What Their New Health Venture Really Means' (*Fortune*, 31 January 2018) <http://fortune.com/2018/01/31/amazon-jpmorgan-berskshire-healthcare/> Last accessed on 1 August 2018.

D. Crow, "IBM strikes digital health deal with Apple, Medtronic and J&J.", *Financial Times*. <www.ft.com/cms/s/0> Last accessed on 3 March 2018.

D.J Armstrong, G.K Meenagh, I. Bickle, et al. Clin Rheumatol, "Vitamin D deficiency is associated with anxiety and depression in fibromyalgia" (2007) 26: 551. <https://doi.org/10.1007/s10067-006-0348-5> Last accessed on 1 June 2018.

Edd Gent, 'AI Is Easy to Fool—Why That Needs to Change', (Singularity Hub, 10 October 2017), <https://singularityhub.com/2017/10/10/ai-is-easy-to-fool-why-that-needs-to-change/> Last accessed on 31 March 2018.

Erika Fry and Sy Mukherjee, 'Tech's Next Big Wave: Big Data Meets Biology' (*Fortune*, 19 March 2018) <http://fortune.com/2018/03/19/big-data-digital-health-tech/> Last accessed on 28 March 2018.

Eugene Kim and Christina Farr, 'Amazon has a secret health tech team called 1492 working on medical records, virtual doc visits', CNBC <https://www.cnbc.com/2017/07/26/amazon-1492-secret-health-tech-project.html> Last accessed on 2 February 2018.

European Commission website, Data Protection < https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/adequacy-protection-personal-data-non-eu-countries_en> Last accessed on 20 May 2018.

Fortune 500, <http://fortune.com/global500/apple/> Last accessed on 2 February 2018.

Fortune Editors and Reuters, 'Amazon, Berkshire Hathaway and J.P. Morgan Are Forming a Non-Profit Health Care Venture' (*Fortune*, 30 January 2018), <http://fortune.com/2018/01/30/amazon-berkshire-hathaway-jpmorgan-nonprofit-healthcare/> Last accessed on 4 February 2018.

Fortune Editors and Reuters, 'Amazon, Berkshire Hathaway and J.P. Morgan Are Forming a Non-Profit Health Care Venture' (*Fortune*, 30 January 2018), <http://fortune.com/2018/01/30/amazon-berkshire-hathaway-jpmorgan-nonprofit-healthcare/> Last accessed on 4 February 2018.

Gartner IT glossary 'Big Data' <http://www.gartner.com/it-glossary/big-data> Last accessed on 31 January 2018.

Natasha Bach, 'First Microsoft, Now Alphabet. Amazon Passes Another Giant to Become The Second Most Valuable U.S. Company', (*Fortune*, 21 March 2018) < http://fortune.com/2018/03/21/amazon-second-most-valuable-company-after-apple/> Last accessed on 30 March 2018.

Natasha Lomas, 'Audit of NHS Trust's app project with DeepMind raises more questions than it answers', (*TechCrunch*, 13 June 2018) < https://techcrunch.com/2018/06/13/audit-of-nhs-trusts-app-project-with-deepmind-raises-more-questions-than-it-answers/?guccounter=1> Last accessed on 19 July 2018.

Nick Triggle, 'Care.data: How did it go so wrong?' *BBC news*, (19 February 2014), <http://www.bbc.com/news/health-26259101> Last accessed on 25 March 2018.

Rachel Z. Arndt, 'Apple is officially in the EHR business. Now what?' (*Modern Healthcare*, 26 January 2018), <http://www.modernhealthcare.com/article/20180126/NEWS/180129910> Last accessed on 2 February 2018.

Randeep Ramesh, 'NHS disregards patient requests to opt out of sharing medical records', (*The Guardian*, 22 January 2015) <https://www.theguardian.com/society/2015/jan/22/nhs-disregards-patients-requests-sharing-medical-records> Last accessed on 3 March 2018.

Samantha Liss, 'Amazon gains wholesale pharmacy licenses in multiple states', (*St. Louis Post-Dispatch*, 27 October 2017), < http://www.stltoday.com/business/local/amazon-gains-wholesale-pharmacy-licenses-in-multiple-states/article_4e77a39f-e644-5c22-b5e6-e613a9ed2512.html> Last accessed on 25 March 2018.

Sejuti Banerjea, 'Amazon on the Brink of EHR Deal with Cerner', *Nasdaq*, (29 November, 2017), <http://www.nasdaq.com/article/amazon-on-the-brink-of-ehr-deal-with-cerner-cm884382> Last accessed on 2 February 2018.

Stephanie Baum, 'Three perspectives on how Amazon could disrupt the pharmacy space', (*MedCity News*, 2 June 2017), <https://medcitynews.com/2017/06/three-perspectives-amazon-disrupt-pharmacy-space/?rf=1/> Last accessed on 4 December 2017.

Other sources:

'FDA approves pill with sensor that digitally tracks if patients have ingested their medication: New tool for patients taking Abilify' USA Food and Drug Administration website, (13 November 2017) <https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm584933.htm> Last accessed on 29 April 2018.

'Survey sheds light on UK's levels of confidence in secure protection of health data', British Journal of Healthcare Computing, (22 May 2017) <http://www.bj-hc.co.uk/publics-trust-protection-health-records-reaches-worrying-levels> Last accessed on 10 August 2018.

Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques (European Commission Working Paper No. 216, 0829/14/EN, 2014).

Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation Techniques (European Commission Working Paper No. 216, 0829/14/EN, 2014).

Article 29 Working Party, Opinion 03/2013 on purpose limitation WP 203 (02.04.2013).

Article 29 Working Party, Working Document 01/2012 on epSOS, WP 189 adopted on 25 January 2012.

Article 29 Working Party, Working Paper nr 131, "Working Document on the processing of personal data relating to health in electronic health records (EHR)", adopted on 15 February 2007.

Cass R. Sunstein and Richard Zeckhauser, 'Overreaction to Fearsome Risks', (2008) HKS Faculty Research Working Paper Series.

Discussion before the European Union Parliament on 9 and 10 October 2012, minutes of the meetings, <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=%2F%2FEP%2F%2FNONSGML%2BCOMPARL%2BPE-504.214%2B01%2BDOC%2BPDF%2BV0%2F%2FEN> Last accessed on 20 March 2018.

Eric Verheul, Bart Jacobs, Carlo Meijer, Mireille Hildebrandt, Joeri de Ruiter, "Polymorphic Encryption and Pseudonymisation for Personalised Healthcare: A Whitepaper, Version 1.1" (Institute for Computing and Information Sciences, Radboud University Nijmegen, The Netherlands).

European Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, 'eHealth Action Plan 2012-2020 - Innovative healthcare for the 21st century', (Communication) COM (2012) 0736 final <http://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:52012DC0736> Last accessed on 4 December 2017.

European Commission, 'Public Consultation on Health and Care in the Digital Single Market', (Strategy on Digital Market Policies – Consultation) <https://ec.europa.eu/digital-single-market/en/news/public-consultation-health-and-care-digital-single-market> Last accessed on 4 December 2017.

Ipsos Mori Research for The Royal Statistics Society. "New research finds data trust deficit with lessons for policymakers." Ipsos MORI. < https://www.ipsos.com/ipsos-mori/en-uk/new-research-finds-data-trust-deficit-lessons-policymakers> Last accessed on 23 March 2017.

J. Oderkirk, 'Readiness of electronic health record systems to contribute to national health information and research', (2017), OECD Health Working Papers, No. 99, OECD Publishing, Paris, <http://dx.doi.org/10.1787/9e296bf3-en> Last accessed on 4 July 2018.

N. Mathers, G. Watt, N. Perrin, 'Towards consensus for best practice: use of patient records from general practice for research', (Welcome Trust, 2009) <https://wellcome.ac.uk/sites/default/files/wtx055661_0.pdf> Last accessed on 1 April 2018.

National Institute for Health and Care Excellence (NICE) Citizens Council, 'What Ethical and Practical Issues Need to Be Considered in the Use of Anonymised Information Derived from Personal Care Records as Part of the Evaluation of Treatments and Delivery of Care?', Citizens Council Reports No. 18, (11 November 2015), 32. <https://www.ncbi.nlm.nih.gov/books/NBK401705/> Last accessed on 28 March 2018.

NHS England, Statement from NHS England and the Health and Social Care Information Centre in response to the Daily Telegraph article, 'Tesco can see your medical records' (10 August, 2015), <https://www.england.nhs.uk/2015/08/response-dt-article/> Last accessed on 25 March 2018.

NHS England. The care.data programme. <https://www.england.nhs.uk/ourwork/tsd/care-data> Last accessed on 25 March 2018.

Nuffield Council on Bioethics, *The Collection, Linking And Use Of Data In Biomedical Research And Health Care: Ethical Issues* (2015) <http://nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data_web.pdf> Last accessed on 16 February 2018.

Office of the Privacy Commissioner of Canada, "Privacy Commissioner releases his finding on the prescribing patterns of doctors", PIPEDA Case Summary #2001-15, Ottawa, October 2, 2001 <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2001/wn_011002/> Last accessed on 30 March 2018.

Oliver Ritchie, Sophie Reid and Lucy Smith (2015) 'Review of public and professional attitudes towards confidentiality of healthcare data', 31. < http://www.gmc-uk.org/Review_of_Public_and_Professional_attitudes_towards_confidentiality_of_Healthcare_data.pdf_62449249.pdf> Last accessed on 1 April 2018.

Privacy International Makes Recommendations To Strengthen UK Data Protection Bill, 1 October 2017 <https://privacyinternational.org/press-release/626/privacy-international-makes-recommendations-strengthen-uk-data-protection-bill> Last accessed on 4 June 2018.

Schippers en Kamp tekenen brede Health Deal voor gerichte beslissingen in de kankerzorg, Rijksoverheid, 8 June 2016 < https://www.rijksoverheid.nl/actueel/nieuws/2016/06/08/schippers-en-kamp-tekenen-brede-health-deal-voor-gerichte-beslissingen-in-de-kankerzorg> Last accessed on 6 June 2018.

Speech by Michel Barnier at the 28th Congress of the International Federation for European Law (FIDE), Lisbon, 26 May 2018, European Commission Press Release <http://europa.eu/rapid/press-release_SPEECH-18-3962_en.htm> Last accessed on 3 June 2018.

The legal framework and guidance on data protection under the Cross-border eHealth Information Services (CBeHIS) T6.2 JAseHN (draft v2 20/10/2016), eHealth Network < https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20161121_co18_en.pdf> Last accessed on 20 May 018.

The White House, 'Consumer Data Privacy in a Networked World: a Framework for protecting privacy and promoting innovation in the global digital economy', (February 2012), <https://www.whitehouse.gov/sites/default/files/privacy-final.pdf>.

UK House of Commons Hansard, 22 March 2018, <https://hansard.parliament.uk/Commons/2018-03-22/debates/ACC7E864-F2E5-4766-8590-FF26CD6C4BB3/LeavingTheEUDataProtectionAgreements> Last accessed on 21 May 2018.

UK Information Commissioner's Office, Letter dated 3 July 2017 to Sir David Sloman, Chief Executive of Royal Free NHS Foundation Trust <https://ico.org.uk/media/action-weve-taken/undertakings/2014353/undertaking-cover-letter-revised-04072017-to-first-person.pdf> Last accessed on 4 August 2018.

UK Parliament website, Bill stages — Data Protection Bill [HL] 2017-19, <https://services.parliament.uk/Bills/2017-19/dataprotection/stages.html> Last accessed on 22 May 2018.

US Health Insurance Portability and Accountability Act (HIPAA).

World Medical Association's Declaration of Geneva, as amended by the 68th WMA General Assembly, Chicago, United States, October 2017. < https://www.wma.net/policies-post/wma-declaration-of-geneva/> Last accessed on 20 March 2018.