ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/357737409

The Right Not to be Subject to Automated Individual Decision Making Profiling Concerning Big Health Data Developing an Algorithmic Culture

Article · November 2021

DOI: 10.51483/IJDSBDA.1.3.2021.6-14

CITATIONS
0

READS
15

3 authors, including:

Yannis Iglezakis
Aristotle University of Thessaloniki

38 PUBLICATIONS **162** CITATIONS

SEE PROFILE

Theodoros Trokanas
Open University of Cyprus

16 PUBLICATIONS **58** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Legal issues concerning the collection and processing of Big Health Data in view of the EU Regulation 679/2016 (General Data Protection Regulation) View project

Project INTERNATIONAL CONFERENCE: NEW TECHNOLOGIES IN HEALTH: MEDICAL, LEGAL AND ETHICAL ISSUES (THESSALONIKI, 21-22 NOVEMBER 2019) View project

All content following this page was uploaded by Theodoros Trokanas on 11 January 2022.

The user has requested enhancement of the downloaded file.

SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Data Science and Big Data Analytics
Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

The Right Not to be Subject to Automated Individual Decision-Making/ Profiling Concerning Big Health Data. Developing an Algorithmic Culture

Ioannis Iglezakis[1*], Theodoros Trokanas[2], Panagiota Kiortsi[3]

[1]Aristotle University of Thessaloniki, Thessaloniki 541 24, Greece. E-mail: iglezakis@hotmail.com
[2]Aristotle University of Thessaloniki, Thessaloniki 541 24, Greece. Email: trokanas@mycosmos.gr
[3] Kapodistrian University of Anthens, Greece. Email: kiortsip@gmail.com

Article Info

Volume 1, Issue 3, November 2021
Received : 27 July 2021
Accepted : 17 October 2021
Published : 05 November 2021
doi: 10.51483/IJDSBDA.1.3.2021.6-14

Abstract

This paper examines the legal issues arising from the collection and processing of Health Big Data in the light of the European legislation for the protection of personal data of natural persons, placing emphasis on the General Data Protection Regulation, i.e., Regulation 679/ 2016. Whether Big Health Data are "personal data" or not is really the heart of the matter. The legal ambiguity is compounded by the fact that, even though the processing of Big Health Data is premised on the de-identification of the data subject, the possibility of a combination of Big Health Data with other data circulating freely on the web or from other data files cannot be excluded. Moreover, data subject's rights, e.g., the right not to be subject to a decision based solely on automated processing, are heavily impacted by the use of AI, algorithms and technologies that reclaim health data for further use, resulting in sometimes ambiguous results that have substantial impact on individuals. On the other hand, as the Covid-19 pandemic has revealed, Big Data analytics can offer crucial source of information. In this respect, this paper identifies and applies the legal provisions to algorithms used in big health data processing, providing an interpretation to address risks to data subject's rights, while embracing the opportunities that Big Health Data has to offer.

Keywords: *Big data, Health data, Genetic data, Privacy, GDPR, Processing principles, Algorithm accountability, Automated profiling*

© 2021 Ioannis Iglezakis *et al.*. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. Introduction

Big Health and/or Genetic Data are a key component of medical evolution. They have considerable potential to improve treatment outcomes, they offer valuable insights, and they improve the quality of a patient's life, in general. In our time, following the Covid-19 outbreak, Big Health Data can offer a valuable arsenal to combat the pandemic using data analytics.

At European level, a first attempt to describe the term "Big Data" was made in 2003. Article 29 Data Protection Working Group in its Opinion on purpose limitation referred to gigantic digital datasets held by corporations, governments and other large organizations, which are then extensively analyzed with computer algorithms. The WP29 also focused on the fact that Big Data can be used to identify more general trends and correlations, but it can

* *Corresponding author: Ioannis Inglezakis, Aristotle University of Thessaloniki, Thessaloniki 541 24, Greece.*
E-mail: iglezakis@hotmail.com

2710-2599/© 2021. Ioannis Iglezakis *et al*. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

also be processed to directly affect individuals (Article 29 Data Protection Working Party, 2013). In 2017 a European Parliament Report defined Big Data as "the collection, analysis and the recurring accumulation of large amounts of data, including personal data, from a variety of sources, which are subject to automatic processing by computer algorithms and advanced data-processing techniques using both stored and streamed data, in order to generate certain correlations, trends and patterns" (European Parliament, 2017). At national level, the Greek National Bioethics Commission deems Big Data as the collection of big and complicated data sets, whose analysis leads to correlations and conclusions that would be unattainable with an analysis restricted to individual data (Hellenic Republic National Bioethics Commission, 2017).

Big data are defined by the "4Vs +1": Volume, Velocity, Variety, Veracity and Value (Hellenic Republic National Bioethics Commission, 2017). In other words, they are characterized by the large volume of information (whether structured or not), the high velocity at which the data are collected and analyzed, the increased complexity and variety of data arriving at different formats, their reliability and the worth produced by their analysis. It is interesting to note that the term 'Big' data is generally understood both in quantitative and procedural terms: it denotes the electronic size of datasets and simultaneously the big computational or human effort it takes to analyze them (Mittelstadt *et al.*, 2016). Moreover, the term 'big' is dynamic in character, as it is conditioned by the advancement level of computing technologies (Mittelstadt *et al.*, 2016). In other words, what is characterized as 'big' today might not to be so in one year or in a decade (Mittelstadt *et al.*, 2016). Finally, as regards the subcategory of genetic data it is correctly underlined that Genomics is an inherently big data science (Mittelstadt *et al.*, 2016). A quintessential example is that even the genome sequence of one single person could be considered Big Data (Stephens *et al.*, 2015).

Big Health and/or Genetic Data sets are multi-sourced, coming *inter alia* from social network sites, hospital records, patient files, and remote sensor networks. The collected data can be structured, unstructured or interrelated, and may include errors or be incomplete. Applications that are based on Internet of Things (IoT) and Artificial Intelligence (AI) process big health data by using machine learning algorithms that perform complex data analytics. Algorithms are mathematical constructs used to turn data into evidence for a given outcome, which is then used to trigger and motivate an action with ethical consequences (Pormeister, 2017). By combining patterns, correlating and adjusting parameters, IoT and AI can be trained to produce new datasets, the so-called 'synthetic data' or 'algorithmically generated data' (Tsamados *et al.*, 2021).

2. Benefits and Risks of Big Health Data

The automated processing of Big Health Data in hospitals dates back to the 1970s: the management, collection and storage of millions of data concerning the medical records of patients, medical examinations and hospital reports was becoming increasingly more time-consuming and costly as health data was expanding. At that time, computer technology was only used to store health data.

Nowadays, AI finds practical application in the prognosis of diseases and recommendations in treatments. The use of Big Health Data analytics, AI, IoT and varying algorithms is well incorporated in almost every healthcare organization at a national and international level. Big Health Data are processed both by the public and private sector health entities, such as, insurance and pharmaceutical companies.

The integration of various data sources means that health data come in many different forms: medical images, texts, numbers, and videos. This "patchwork" of different data will be processed with the application of algorithmic techniques to generate new set of data that identify and correlate new patterns (Tsamados *et al.*, 2021). For example, AI tools track patterns in Big Health Data that can be used for image analysis. Machine learning is then able to identify types of cancer and Covid-19 from images that depict infected organs of patients. Similarly, AI technology applications can predict when certain health incidents, such as, an ischemic stroke or hemorrhage, are bound to occur to trauma patients (Yeung, 2019). Further, AI and IoT technologies have revolutionized patient care by helping patients and elderly persons in their daily lives at home and hospitals by offering immediate intervention. Wearable devices (Price and Cohen, 2019) and robots can provide nursing care without the danger of transmitting diseases (European Economic and Social Committee, 2017).

Finally, the integration of many sources of medical data proved indispensable for the deployment of Big Data in medical research like the monitoring of therapeutic protocols and medical methodologies in order to study the prevention and treatment of various forms of diseases. Ever since, Big Data have been a fundamental part of medical research as the latter presumes upon clinical trials' results, meta-analysis and systematic review (Arwidson *et al.* 2017).

On the other end of the spectrum, Big Health Data processing involves risks for the rights and freedoms of data subjects. As mentioned in the WP29 'Guidelines on Automated individual decision-making and profiling for the purposes of Regulation 2016/679', EU policymakers grew deeply concerned with the risks emanating from the use of Big Data analytics tools and algorithms in automated decision-making and profiling; namely, the ethical implications in terms of fairness, transparency, and accountability.

Automated decision-making is the ability to make decisions by technological means without human involvement (Dash *et al.*, 2019). According to the provisions of Article 4 (4) GDPR and Article 3 (4) of Directive 2016/680 as well as Article 4 (5) of Regulation 2018/1725 profiling is the automated processing of personal data with a view to evaluating, i.e., analyzing or predicting personal aspects of a natural person, such as natural work performance, economic situation, health, personal preferences, interests, reliability, behavior, location or movements. Automated decision-making and profiling are not mutually exclusive activities: automated decisions can be taken with or without profiling; conversely, profiling can take place without making automated decisions (Article 29 Data Protection Working Party, 2018). Besides, something that begins as a simple automated decision-making process could turn into one based on profiling (Article 29 Data Protection Working Party, 2018). Unsurprisingly, the drafters of the EU law on personal data protection (Article 29 Data Protection Working Party, 2018) expressed major concerns about the potential risks involved for the interests and rights of the data subject as well as about discriminatory effects on natural persons on the basis of special category data due to automated individual decision-making and profiling.

Automated decisions can be based on any type of data (Recital 71 and Recital 43). In fact, apart from personal data of the person targeted by the decision, they may include personal data of third person, and even non-personal data (Article 29 Data Protection Working Party, 2018). On the contrary, decisions based on profiling must involve automated processing of personal data, as it follows from the definition of the term 'profiling' in the provisions of Article 4 (4) GDPR.

Even though the use of automated processing and/or profiling allows for fastest and most cost-effective applications, it certainly entails some shortcomings.

First of all, massive algorithmic models and datasets are characterized by a technical complexity, which lies beyond human cognitive ability (Bygrave *et al.*, 2020). To put it simply, humans cannot understand the process by which a computer receives data, processes them and responds by analyzing them. This limited understanding is rooted in either the inability of data subjects to comprehend technical aspects of the procedure or the malleability of algorithms, which are continuously reprogrammed (Tsamados*et al.*, 2021). Above all, this lack of explainability is the main source of a lack of transparency, which proves inherent in machine learning algorithms (Tsamados*et al.*, 2021).

In addition, the validity of automated decision making and/or profiling is highly questionable. Machine learning algorithms may produce inconclusive results, because either they are fed with poor quality data or they focus on mere associations and correlations instead of causal connections (Tsamados*et al.*, 2021). For instance, it has been reported that the accuracy of algorithms designed to predict patients' prognoses in clinical settings may be undermined because of the total reliance of algorithms on quantifiable data inputs (i.e., vital signs, previous success rates of comparative treatments) and of the ignorance of emotional factors (Tsamados *et al.*, 2021). Also, the exclusion of certain data variables or contexts may produce inaccurate or systematically biased algorithmic decisions, leading to unfair discrimination (Tsamados *et al.*, 2021) (i.e., underrepresentation or misrepresentation of population groups), while in other cases it might be challenging to monitor the performance of algorithm analytics in order to evaluate its efficiency (Tsamados *et al.*, 2021). For example, healthcare resource allocation decisions generated by algorithms may be flawed if a research hospital healthcare algorithm is applied to a rural clinic on the assumption that both the research hospital and the rural clinic enjoy the same level of resource availability (Veale and Brass, 2019). In a nutshell, when data controllers use erroneous data as a basis for developing a problem-solving machine learning process, ambiguous, biased, or misguided solutions will arise with substantial impact on individuals' lives (Tsamados *et al.*, 2021). After all, as far as algorithms are concerned, one size does not fit all.

Furthermore, some services providers, such as Facebook and Google, grant access to third party websites without clearly and efficiently informing data subjects about which data will be transferred to third parties and who will store them. Similar omissions occur with information regarding the recipients' identity, the period of processing and the purpose for which the data will be used or re-used (Kaltheuner *et al.*, 2018). One of the most conspicuous examples was the black box practice in the Cambridge Analytica case (European Economic and Social Committee, 2017), in which it was revealed that data subjects (Facebook users) had been deceived regarding data recipients and data processing purposes.

Admittedly, the biggest challenge is that Big Health Data can become re-identifiable. In other words, health data are unique by definition, hence the initial data subject can be traced really easily (Isaak and Hanna, 2018).

Re-identification can occur as the result of algorithms tracking repetitive patterns (Price and Cohen, 2019). Specifically, breaking patterns or cases that lie far from the majority can be related to a specific person, distinguishing it from the anonymous crowd. Consequently, the concept of "unidentifiable" data is fluid. As a matter of fact, it is technological developments that dictate what is unidentifiable.

3. The Regulation of the Right not to be Subject to Automated Individual Decision-Making or Profiling

In this section it will be discussed whether Big Health Data analytics, AI, IoT and machine learning algorithms fall within the scope of the EU personal data protection legislation.

Building on Article 15 of the repealed Directive 95/46, the provisions of Article 22 of Regulation 2016/679 and Article 11 of Directive 2016/680 address the issue of automated individual decision-making processing and profiling. In more detail, the application of Article 22 (1) of Regulation 2016/679 presupposes three elements: (i) a decision; (ii) based solely on automated decision or profiling; and (iii) producing legal effects concerning data subject or significantly affecting him or her (Criado and Such, 2019). As far as the wording of these provisions is concerned, both Regulations refer to a data subject's right not to be subject to a decision taken based solely on automated processing, including profiling (According to Recital 71 of Regulation 2016/679), whereas the Directive sets forth a corresponding prohibition. The different approach between the Regulations and the Directive implies a legal divide over whether we are dealing here with a right, as suggested by the wording of Article 22 GDPR or with a prohibition, as suggested by the wording of Article 11 of Directive 2016/680 (The same right is proclaimed in Article 9 (1) (a)).

In principle, none of the said provisions endorses automated individual decision-making processing and profiling. Nonetheless, Article 22 (2) of Regulation 2016/679 and Article 11 (1) of Directive 2016/680 introduce a first set of qualifications, which unlock such processing. In more detail, in the two Regulations the existence of (a) a contract between the data subject and the data controller; or (b) statutory authority (i.e., European or Member State law authorization); or (c) the explicit consent of the data subject suffices to justify such processing, while in the Directive only the existence of statutory authority legitimizes criminal offence and criminal penalty data processing. It is worth clarifying that such circumstances function irrespective of their being regarded as legal restrictions to a right or as reasons for lifting a prohibition. Additionally, even if a contract (a) or explicit consent (c) exists, Article 22 (3) of Regulation 2016/679 obliges the data controller to implement suitable measures to shield the data subject's rights and freedoms and legitimate interests, setting as a threshold its right to request the controller's human intervention, to express his or her viewpoint and to challenge the decision.

According to the provisions of Article 22 (4) of Regulation 2016/679 the decisions referred to in Article 22 (2) of Regulation 2016/679 should not to be based on special categories of data. To put it differently, it seems that, in principle, automated individual decision-making processing and profiling based on special categories of data are prohibited.

Nevertheless, according to the said articles even this prohibition is lifted if two of the general derogations of Article 9 (2) (a) or (g) of Regulation 2016/679 and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place. A systematic interpretation of Articles 22 (4) and 22 (2) of Regulation 2016/679 leads us to assume that automated decision-making and profiling involving special categories of personal data is allowed only if the following cumulative conditions exist: a specific exemption of Article 22 (2) of Regulation 2016/679 combined with either of the general derogations (a) or (g) of Article 9(2) of Regulation 2016/679 (Christopher *et al.*, 2020).

It is reminded that Big Health Data and Big Genetic Data fall under the exhaustive list of special categories of personal data of Article 9 of Regulation 2016/679 . Consequently, automated decision-making and profiling based on Big Health (or Genetic) Data could potentially take place in three alternative legal contexts: (i) a contract between the data subject and a data controller; (ii) European or Member State law authorization; and (iii) a data subject explicit consent, in conjunction with either the data subject explicit consent or the existence of a substantial public interest based on European or Member State law. Apart from the apparent overlap between the two sets of legal contexts as far as the explicit consent is concerned, one might wonder which of them can legitimize Big Health (or Genetic) Data processing.

On one hand, it is well-known that Big Health (or Genetic) Data collection usually remains unnoticed, and the tools and techniques of their processing are unclear. On the other hand, as a rule, their collection and processing go beyond the primary purpose, involving a diversion of the initial purpose of collection and further processing for a secondary purpose (Article 29 Data Protection Working Party, 2018). Moreover, the processing of non-sensitive personal data, such as, personality traits, interests, financial situation, and other facts, can result to sensitive data (i.e., health data) depending on the AI methods and algorithms that are applied at the processing (Hellenic Republic National Bioethics Commission) As a result, it is questionable whether the rights and legitimate interests of data subjects will be preserved, since there is no way to predict the outcome, e.g., the category of data that will be generated, in order to apply the legal provisions regarding the processing of special categories of data.

Against this backdrop, it is extremely unlikely for a data subject to have given an explicit consent to automated decision-making and profiling processing based on Big Health (or Genetic) Data, let alone to have entered into a contract with the data controller for such processing. This leaves us with no option but to rely on EU or Member State law which would authorize decisions based on automated processing of Big Health Data, according to Article 22 para. 2 lit. b of Regulation 2016/679, provided that suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are laid down in the EU or national law. To conclude, automated individual decision-making processing and profiling based on Big Health and/or Genetic Data cannot be lawfully performed, unless relevant European or Member State laws are enacted.

If EU or Member State laws are enacted to regulate automated individual decision-making processing and profiling based on Big Health and/or Genetic Data, they must be premised on a set of guiding principles: respect for the right to information of the data subject as well as for the fairness, transparency and accountability of processing.

4. The Legal Requirement for Algorithmic Information, Fairness and Transparency

Apart from Article 22 of Regulation 2016/679, the provisions of Articles 13 and 14 of Regulation 2016/679, which specify the right of the data subject to be informed in the framework of automated decision-making and profiling, also apply to Big Health Data processing. All the above-mentioned Articles stipulate that the controller shall share with the data subject purposeful information about the logic involved, as well as the significance and the possible repercussions of such processing for the data subject.

To be more specific, WP29 Guidelines (Kaltheuner and Bietti, 2018) expect the controller to explain the rationale behind or the criteria relied on in reaching the automated decision and the information to be sufficiently comprehensive for the data subject, so that he/she understands the reasons for the automated decision (Article 29 Data Protection Working Party, 2018). By contrast, the controller is under no obligation to give a complex explanation of the algorithms used or to disclose the full algorithm (Article 29 Data Protection Working Party, 2018). To elaborate on the scope of the right to information, W29 enumerates some good practice recommendations: the controller should inform the data subject about the data categories used or to be used in the profiling or decision-making process; he should explain the relevance of the data, the construction of any profile used in the automated decision-making process, including any statistics used in the analysis, the relevance and the use of this profile to the automated decision-making process (Article 29 Data Protection Working Party, 2018).

Inescapably, by meeting the legal requirements of the right of the data subject to be informed in the context of automated decision-making and profiling, a controller also adheres to the fundamental principles of algorithmic fairness and algorithmic transparency, which are enshrined in Articles 5 (1) (a) of Regulation 2016/679. As Article 29 Data Protection Working Party points out, transparency is paramount, because profiling may lead to unfair discrimination (Article 29 Data Protection Working Party, 2018). As far as big health data are concerned the danger lies into discriminating against “genetically inferior” people or denying insurance.

5. The Legal Requirement for Algorithmic Accountability

Contrary to Directive 95/46, which did not explicitly mentioned the term accountability, Article 5 (2) Article 29 Data Protection Working Party, 2018 and Article 4 (4) of Directive 2016/680 introduce a general principle of accountability, whereby the controller shall be responsible for, and be able to demonstrate compliance with the basic principles relating to processing of personal data (OECD, 1980) . In other words, the aforementioned provisions impose the responsibility for the compliance of processing with the European data protection legislation and the burden of proof for such compliance onto the controller (Voigt, 2017). What is essential here is that accountability is not a standalone principle, but it is substantiated by the other obligations of the data controller under the European data protection legislation, as it is suggested by the direct references of Article 5 (2) to Article 5 (1) of Regulation 2016/679 and of Article 4 (4) to Article 4 (1), (2) and (3) of Directive 2016/680 (Voigt, 2017). Broadly speaking, the accountability principle acts as a counterbalance to the increased power of the data controller towards the data subject, which personal data processing entails (Alhadeff *et al.*, 2012).

In practice, the accountability principle is understood to convey two elements: (i) the requirement for the data controller to take appropriate and effective measures to implement data protection principles; (ii) the requirement to demonstrate upon request the appropriate and effective measures taken (Article 29 Data Protection Working Party, 2010).

Given the deluge of Big Data processing in our era the notion of accountability had to be redefined. To this end, a Report published in 2017 (European Parliament, 2017) by the European Parliament coined the terms "algorithmic accountability" and "algorithmic transparency" to represent a pair of basic principles of personal data processing applied to a Big Data framework.

Unsurprisingly, the same approach was endorsed by some legislators. One prime example is the US Algorithmic Accountability Act of 2019 (https://www.congress.gov/bill/116th-congress/house-bill/2231/text), which was introduced in House on November 4, 2019, but has yet to be enacted. The Bill mandates commercial entities which meet certain criteria (Section 5 (B)) to conduct data protection impact assessments of high-risk automated decision systems. Such criteria are 50.000.000 $ annual gross income for a 3-year preceding period or possession/control of personal information on more than 1,000,000 consumers or 1,000,000 consumer devices or just being data brokers. Evidently, the proposed US Bill reflects the spirit of the provisions of Article 35 of Regulation 2016/679, Article 39 of Regulation 2018/1725 and Article 27 of Directive 2016/680.

According to Section 2 (1) of the proposed Bill, the term "automated decision system" means a computation, including one coming from machine learning, statistics, or other data processing or artificial intelligence techniques, that helps to reach a decision or facilitates human decision making, that affects consumers. According to Section 2 (7), an automated decision system is high-risk if (a) it poses a significant risk to the privacy or security of personal information of consumers or results in or contributes to inaccurate, unfair, biased, or discriminatory decisions impacting consumers; (b) it makes decisions or facilitates decision-making based on systematic and extensive evaluations of consumers, including attempts to analyze or predict sensitive aspects of their lives, e.g., their health; and (c) it involves the personal information of a significant number of consumers regarding inter alia health and genetic data.

According to Section 2 (2) of the proposed Bill assessments of high-risk automated-decision systems must include at least: (a) a detailed description of the design, the purpose and the data that are processed by the automated decision system; (b) a cost and benefit analysis; (c) a privacy and security risk assessment; and (d) the measures required to mitigate the risks.

6. Concrete Proposals for the Protection of Data Subject Rights

As already mentioned, the provisions of the EU legislation on data protection (See Articles 22 (2) (b), 22 (3) and 22 (4)) provide that the data controller should implement suitable measures to safeguard the rights, freedoms, and legitimate interests of the data subject regarding automated individual decision-making and profiling. The controller should adopt measures that safeguard the data subject's rights such as expressing their opinions and to challenging the decisions that are based on automated processing.

Regrettably, the European data protection laws provide no uniform recommendations about how the controller should implement appropriate measures and safeguards to ensure that the data subject's rights will not be encroached.

On the other hand, the use of algorithms in Big Health and/or Genetic Data needs to ensure that data used are adequate in quality to enable drawing the right conclusions. Most importantly, the Drafters of European data protection legislation insist that the controller implement appropriate, technical, and organizational measures to ensure that factors resulting in inaccuracies in personal data are corrected and the risk of errors is minimized (Recital 71 Regulation 2016/679). In truth, to overcome the problem of collecting and processing unreliable or biased data, various methods have been proposed (Recital 71 Regulation 2016/679).

In addition, since algorithmic tools play a crucial role in Big Health Data processing and pose a serious risk for individual autonomy, risk must be assessed to protect the privacy of data subjects and to combat discrimination (Favaretto *et al.*, 2019). What this means is that the impact assessment regarding high-risk automated-decision systems should be focused on the algorithm (algorithmic impact assessment), transcending a typical DPIA.

As the European Group of Ethics in Science and new Technologies describes (Yeung and Lodge, 2019), the challenges posed using new technologies in Big Health Data processing must be dealt with regulation or governance which is mindful of fundamental individual rights, such as the right to privacy, freedom of research and innovation. To measure compatibility with these fundamental rights a Human Rights Impact Assessment (HRIA) would be a step in the right direction. HRIA is broadly defined as a process of identifying, understanding, assessing and addressing the adverse effects of various activities on the enjoyment of human rights of impacted rights-holders (European Group on Ethics in Science and New Technologies, 2018). Using internationally recognized human rights standards and principles as an authoritative benchmark for the impact assessment, HRIA is considered to constitute an implementation of due diligence obligation (Gotzmann, 2019). Traditionally, HRIA is associated with equality and non-discrimination and it primarily focuses on access to information, transparency and accountability (Gotzmann, 2019). Applying a HRIA on Big Health (or Genetic) Data processing will help shift focus from data protection to ethics (Gotzmann, 2019).

7. Conclusion

Nowadays, Big Health and/or Genetic Data processing increasingly resorts to Artificial Intelligence, the Internet of Things, and machine learning algorithms to facilitate automated decision making or/and profiling. Apparently, this phenomenon is a double-edged sword. Stakeholders can reap benefits, as long as they are prepared to assume increased risks. The European personal data protection legislative trio (Regulation 2016/679, Directive 2016/680, Regulation 2018/1725) establishes the principle of prohibition of automated individual decision-making processing and profiling based on Big Health and/or Genetic Data, a prohibition which is lifted if a combination of conditions is met. As it is extremely unlikely for a data subject to have given an explicit consent to or to have entered into a contract with the data controller for automated decision-making and profiling processing based on Big Health (or Genetic) Data, such processing cannot be lawfully performed, unless relevant European or Member State law authorization exists. The enactment of pertinent European or Member State legislation should feature a reshaping of core rights, principles, and processes of current personal data protection legal instruments; notably, an adjustment of the right to information of data subjects; an upgrade of the general requirements for fairness, transparency and accountability to algorithmic fairness, algorithmic transparency and algorithmic accountability; a possible extension of the scope of a Data Protection Impact Assessment (DPIA) to embrace a Human Rights Impact Assessment (HRIA). In that sense, the proposed US Algorithmic Accountability Act of 2019 could be used as a model for regulating Big Health Data processing.

Funding

This paper is based on the research project “Legal issues concerning the collection and processing of Big Health Data in view of the EU Regulation 679/2016 (General Data Protection Regulation)” (MIS 5047861) of Operational Program “Human Resources, Development, Education and Lifelong Learning” - NSRF 2014-2020, which is co-financed by the ESF and national funds.

References

Alhadeff, J., Van Alsenoy, B., and Dumortier, J. (2012). The Accountability Principle in Data Protection Regulation: Origin, Development and Future Directions, In D. Guagnin, D., Hempel, L., and Ilten, C. (eds.). *Managing Privacy through Accountability*, 54, Palgrave Macmillan.

Article 29 Data Protection Working Party (2010). Opinion 3/2010 on the principle of accountability, adopted on July 13.

Article 29 Data Protection Working Party. (2018). Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation 2016/67, Adopted on October 3, 2017 as last Revised and Adopted on February 6.

Article 22, In C. Kuner, L.A. Bygrave, C. Docksey, L. Drechsler (ed.) (2020). The EU General Data Protection Regulation (GDPR) A Commentary, Oxford University Press, 2020, Under C 3, p 533.

Cohen, I.G., Lynch, H.F., and Vayena, E.U. (2018). Gasser, Introduction, In Cohen, I.G., Lynch, H.F., and Vayena, E.U (eds.). *Big Data, Health Law, and Bioethics*, 8, Cambridge University Press, p. 1.

Convention 108+ of the Council of Europe for the protection of individuals with regard to the processing of personal data, as amended by Protocol CETS No. 223.

Criado, N., and Such, J. (2019). Digital Discrimination in: An Introduction, In Yeung, K., and Lodge, M. (eds.). *Algorithmic Regulation*, Oxford University press, p. 11.

Dash, S., Shakyawar, S.K., Sharma, M. *et al*. (2019). Big Data in Healthcare: Management, Analysis and Future Prospects. *J. Big Data*, *6*, 54, https://doi.org/10.1186/s40537-019-0217-0

Directive 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

European Economic and Social Committee. (2017). The Ethics of Big Data: Balancing Economic Benefits and Ethical Questions of Big Data in the EU Policy Context, p. 33, Under 5.1.2, available at https://www.eesc.europa.eu/sites/default/files/resources/docs/qe-04-17-306-en-n.pdf.

European Group on Ethics in Science and New Technologies. (2018). Statement on Artificial Intelligence, *Robotics and 'Autonomous' Systems*, 15-20, Brussels.

European Parliament (2017). Committee on Civil Liberties, Justice and Home Affairs, Report on Fundamental Rights Implications of Big Data: Privacy, Data Protection, Non-discrimination, Security and law-enforcement (2016/2225(INI)), 20 February 20, Under A.

Evidently, the proposed US Bill Reflects the Spirit of the Provisions of Article 35 of Regulation 2016/679, Article 39 of Regulation 2018/1725, and Article 27 of Directive 2016/680.

Favaretto, M., *et al.* (2019). Big Data and Discrimination: Perils, Promises and Solutions. A Systematic Review. *J Big Data*, 6, available online https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0177-4#citeas

Harrison, J. (2011). Human Rights Measurement: Reflections on the Current Practice and Future Potential of Human Rights Impact Assessment, *Journal of Human Rights Practice*, 3(2), 162-187, available at https://doi.org/10.1093/jhuman/hur011

Hellenic Republic National Bioethics Commission, Opinion, Big Data in Health, p. 5, available at http://www.bioethics.gr/images/pdf/GNOMES/OPINION_Big_Data_FINAL_EN.pdf.

International Standards on the Protection of Personal Data and Privacy, (2009). The Madrid Resolution-Joint Proposal for a Draft of International Standards on the Protection of Privacy, p.13. https://edps.europa.eu/sites/edp/files/publication/09-11-05_madrid_int_standards_en.pdf

Isaak, J., and Hanna, M.J. (2018). User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection, In *Computer, 51*(8), 56-59.

Jørgensen, R. *et al.* (2019). Chapter 12: Exploring the role of HRIA in the information and Communication Technologies (ICT) sector, In Götzmann, N. (ed.), *Handbook on Human Rights Impact Assessment, Research Handbooks on Impact Assessment Series*, 205-218, Edward Elgar.

Kaltheuner, F., and Bietti, E. (2018). Data is Power: Towards Additional Guidance on Profiling and Automated Decision-making in the GDPR, *Journal of Information Rights Policy and Practice*, *2*(2), 8.

Mittelstadt, B.D., and Floridi, L. (2016). Introduction, In Mittelstadt, B.D., and Floridi, L. (Ed.), *The Ethics of Biomedical Data, Law Governance and Technology Series*, 29(Springer), 2.

OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, Recommendation of the Council of Organisation for Economic Co-operation and Development of 23rd September 1980-Guidelines on the Protection of Privacy and Transborder Flows of Personal Data. https://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm#recommendation

Pormeister, K. (2017). The GDPR and Big Data: Leading the Way for Big Genetic Data?, In Schweighofer, E., Leitold H., Mitrakas A., and Rannenberg, K. (Eds). *Privacy Technologies and Policy*. APF. Lecture Notes in Computer Science, *10518*(Springer), 13.

Stephens, Z.D. *et al*. (2015). Big Data: Astronomical or Genomical? *PLoS Biol.*, *13*(7), e1002195, 1.

Tsamados, A., Aggarwal, N., Cowls, J. *et al.* (2021). The Ethics of Algorithms: Key Problems and Solutions. *AI & Soc.*

Veale, M., and Brass, I. (2019). Administration by algorithms Public Management Meets Public Sector Machine Learning in K Yeung, M Lodge (ed.), *Algorithmic Regulation*, 133-138, Oxford University Press.

Voigt, P., and von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR). *A Practical Guide*, 31, Springer.

Yeung, K., and Lodge, M. (2019). Algorithmic Regulation: An Introduction, In Yeung, K., and Lodge, M. (eds.), *Algorithmic Regulation*, 14-16, Oxford University Press.

Cite this article as: Ioannis Iglezakis., Theodoros Trokanas., and Panagiota Kiortsi (2021). The Right Not to be Subject to Automated Individual Decision-Making / Profiling Concerning Big Health Data. Developing an Algorithmic Culture. *International Journal of Data Science and Big Data Analytics*, 1(3), 6-14. doi: 10.51483/IJDSBDA.1.3.2021.6-14.

View publication stats