Routledge
Taylor & Francis Group

The New Bioethics

A Multidisciplinary Journal of Biotechnology and the Body

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ynbi20

Data Ethics in Digital Health and Genomics

Muhammed Erkan Karabekmez

To cite this article: Muhammed Erkan Karabekmez (2021) Data Ethics in Digital Health and Genomics, The New Bioethics, 27:4, 320-333, DOI: 10.1080/20502877.2021.1996965

To link to this article: https://doi.org/10.1080/20502877.2021.1996965

THE NEW BIOETHICS, Vol. 27 No. 4, 2021, 320–333

Data Ethics in Digital Health and Genomics

MUHAMMED ERKAN KARABEKMEZ
Department of Bioengineering, Istanbul Medeniyet University, Istanbul, Turkey

The digital revolution has disruptively reshaped the way health services are provided and how research is conducted. This transformation has produced novel ethical challenges. The digitalization of health records, bioinformatics, molecular medicine, wearable biomedical technologies, biotechnology, and synthetic biology has created new biological data niches. How these data are shared, stored, distributed, and analyzed has created ethical problems regarding privacy, trust, accountability, fairness, and justice. This study investigates issues related to data-sharing permissions, fairness in secondary data distribution, and commercial and political conflicts of interest among individuals, companies, and states. In conclusion, establishing an agency to act as deputy trustee on behalf of individuals is recommended to intermediate the complex nature of informed consent. Focusing on decentralized digital technologies is recommended in order to catalyze the utilization of data and prevent discrimination without circulating data unnecessarily.

KEYWORDS Data ethics, bioethics, digital health, genomics, bioinformatics

Introduction

The immense production and swift transfer of data as well as the ability to store huge amounts of data make up the digital revolution and has shaped today's world in various aspects such as finance, marketing, entertainment, communication, research, and health. The digital revolution has also created novel ethical issues in many fields. The digitalization of health systems, bioinformatics, advancements in genomics, the use of the Internet of things (IoT) in biomedical devices, cloud-based technologies, and the use of artificial intelligence (AI) in medical and biological applications have brought out particularly challenging issues in bioethics.

Computer ethics have been discussed as a term only since 1985 (Moor 1985), while bioethics as a term is no older than 50 years (Martensen 2001). These terminologies were coined by scholars after a groundswell of applied ethics occurred

 DOI 10.1080/20502877.2021.1996965

in the 1970s. Soon after, rapid changes in technology and medical applications introduced even more specific ethical discussions. Starting from the early days of the digital revolution, some efforts like e-Health code of ethics (Rippen and Risk 2000) were proposed to address the emergent issues. Not just in the medical context but also in general, data ethics has been a hot topic in the field of ethics research (Hand 2018).

Conventional paper-based health records being transformed into modern digital health records has brought new discussions around sensitive data privacy and data protection. Patients' data security has been a topic in ethics for a long time (Declaration of Helsinki 1964). In fact, the ethical dimension of digital health records is not so different than that of previous decades; the only things that have changed are the type of breaches and the measures required to be taken against possible data breaches. Ethical principles like privacy and trust can have dynamic and contextual meanings, and a proper combination of these principles can help form a strong digital health system (Ruotsalainen and Blobel 2020). If one assumes the only problem to be unauthorized access to health records whether accidentally or maliciously then prevention can be attained through policies, education, preventive engineering, and deterrence (Myers *et al.* 2008; Lee 2017). However, the problem can become more complicated with the unceasing tension between public concerns and individual rights.

Another important area in data is bioinformatics, which is a data science subdiscipline dealing with the molecular data concerning living cells, including genetic data. As Hongladarom (2006, p. 40) pointed out, the 'ethics of bioinformatics is an amalgam of the two major strands of applied ethics' at the intersection of bioethics and computer ethics. Genetic data is closely linked to individual identity, and its privacy and security in terms of ethics should at least be more crucial than individual financial data, for example. After the completion of the Human Genome Project in 2001, genetic research and the clinical function of genetics changed drastically, becoming more data-centric. This transformation in the field also affected the ethical issues surrounding genetics (UNESCO 2003). Previous issues were more concentrated on genetic counselling, genetic discrimination (EC Oviedo Convention 1997), and genetically modified agricultural products. After this transformation: reciprocity, mutuality, solidarity, citizenry and universality were proposed as new aspects in the ethics of genetics (Knoppers and Chadwick 2005). Early ethical considerations in the genomic era also involved the participation of close family members in the consent process and sharing data procedures with research participants (McGuire *et al.* 2008). USA's Presidential Commission for the Study of Bioethical Issues (PCSBI) recommended five bioethical principles; (1) public beneficence, (2) responsible stewardship, (3) intellectual freedom and responsibility, (4) democratic deliberation, and (5) justice and fairness in addressing technological developments in synthetic biotechnology (PCSBI 2010). It re-recommended the same principles for addressing ethical issues regarding whole genome sequencing (PCSBI 2012).

Anonymizing personal data is the most common way to utilize data for public benefits without violating personal autonomy. Anonymizing genomic data is troublesome; removing all personal identifiers is insufficient because the data itself is

self-identifying and re-identification is possible to some extent. The proposed solution for the difficulty of anonymizing genomic data is to either have elaborated restrictions on data access or informed consent (Wjst 2010).

The future of medicine has been formulated as P4 medicine (predictive, personalized, preventive and participatory) (Auffray *et al.* 2010). Participation occurs through wearable biomedical technologies that enable continuous personal health data to be shared with clinicians. This trend will move health care beyond clinics to the public sphere and also effect the identification of health and disease. Dynamic data will change the conceptualization of disease from a static state to a cascading process (Boenink 2010). P4 medicine (i.e. precision medicine) focuses more on wellbeing than on treating disease. In this context, wearable biomedical devices are crucial for monitoring one's health status in terms of sleep quality, air quality, environmental exposome, diet, biomarker status, body posture, stress status, blood pressure, and physical activities (Shameer *et al.* 2017). Some biomedical devices (i.e. digital medicine) not only monitor patients' data but also treat the patients automatically. a typical example is devices for diabetes that monitor blood-glucose levels and auto-dose insulin injections. Ethical issues regarding biomedical technologies and digital health devices are not so different compared to ethical issues about other contemporary data-related health technologies (e.g. data security, informed consent, and data management policies; Nebeker *et al.* 2017).

Molecular medicine is an emerging sub-discipline that utilizes bioinformatics tools for analyzing big data coming from the molecular states of living cells and environmental exposome in order to achieve the goals of precision medicine. Molecular medicine focuses on discovering biomarkers, and these discoveries are able to reveal potential health risks with some uncertainty. This will also bring inevitable ethical issues about the 'desirability of knowledge of future health risks' (Boenink 2010, p. 18).

New biomedical technologies can also be used in national security. USA's Defense Advanced Research Projects Agency (DARPA) has several projects regarding neuroscience for developing enhanced soldiers, brain-computer interfaces in flight jets, and new interrogation techniques (Tennison and Moreno 2012). Unsurprisingly, serious ethical debates exist around these projects.

Even though not having emerged in any medical or biological context, some digital tools can be adopted by the biology and medical communities. For example, digital storytelling is common in social media and has started to be used with public health purposes. The use of digital storytelling in public health for awareness about health issues brings ethical issues not only regarding the use of digital outputs but also the way they are produced (Gubrium *et al.* 2014).

The synthetic biology equipped with digital tools in this new era has also produced novel ethical problems in biotechnology. The J. Craig Venter Institute announced in 2008 that they had synthesized the first human-made artificial bacterial genome of about 583 kb (J. Craig Venter Institute 2008). Soon after, the same institute announced that they had achieved replicating a synthetic bacterial genome in a living bacterial cell (J. Craig Venter Institute 2010). As the founder of the institute, J. Craig Venter made sure to write a genetic code on computers,

implement the code in the lab using certain chemical mixtures, and achieve the creation of life (Venter 2014). Some people started to accuse Craig Venter of playing god because of his bold assertions (Peters 2018), while others defended his position and critiqued those accusations for being unsecular (van den Belt 2009).

The most discussed topic of ethics in the digital age has probably been AI (Montreal AI Ethics Institute 2021). AI has applications in all possible biological and medical fields such as research on diagnostic and prognostic tools (Panch *et al.* 2018), pattern recognition in radiology (Pesapane *et al.* 2018) and histopathology (Colling *et al.* 2019), molecular medicine, medical decision support systems, and biotechnology (Shah *et al.* 2019).

This paper investigates ethical problems regarding data in biological and health systems through six aspects; (i) complications around individual permissions for accessing data, (ii) the commercial and political aspects of secondary data sharing, (iii) biotechnology and synthetic biology, (iv) biomedical devices, (v) genetic surveillance, and (vi) AI. The conclusion section recommends and summarizes possible solutions.

Allowing individual access to data

The rapid pace of technology has been able to bring new methods such as encryption algorithms, cloud solutions, and digital information management systems for preventing digital data breaches (Myers *et al.* 2008). However, each new technology comes with its own backdoor, with digital piracy developing at the same pace.

In light of the progress of the past decades, more holistic approaches like systems biology have replaced conventional gene-centric molecular biology. Emerging perspectives have highlighted the crucial role of other molecular entities apart from DNA, such as proteins and RNA. The United Nations Educational, Scientific, and Cultural Organization (UNESCO) has accordingly addressed genetic data alongside proteomic data in its declarations (UNESCO 2003). Thus, one may even consider extending the perception of genetic data to cover data regarding all molecular entities in living cells. Genomic data have huge potential in biotechnological and biomedical research and creates another dimension in the tension between public interests and personal autonomy. The obstacle to anonymizing personal genomic data due to the potential risk of re-identification has led to more stringent measures in accessing data. This has created the limited utilization of genomic data in research, which has also been able to bring potential barriers to research topics like drug development. Early ethical evaluations of genomics suggested that only the clinically relevant summary of genomic data should be shared with physicians (McGuire *et al.* 2008) or researchers. This could be useful under limited conditions. For instance the Cancer Genome Atlas Program (TCGA) under the National Cancer Institute and National Institute of Health publicly provides RNAseq datasets from tumour samples after certain data preprocessing steps (TCGA 2021). In its raw form, the RNAseq data is complementary to the DNA data; that's why it is quite sensitive. After being processed into a quantitative dataset (i.e. 'counts'), it becomes more suitable for anonymization. However, the progress in the field

does not always match that limitation. A clinical evaluation of genomic data is not straight forward, which makes defining something as clinically relevant difficult, consequently clinicians have become more and more involved in bioinformatics analysis directly on their own (Long and Alpern 2009). Therefore, limiting physicians' access to genomic data is difficult, even if the individual is favoured over the public.

Data sharing can be even more crucial in some research contexts such as rare diseases where the number of patients is limited number and no easy way exists to conduct research on the cause or progress of these kinds of diseases without sharing data. Some digital platforms have attempted to reach out to patients with rare disease and match them using their genetic and phenotypic profiles through a process of informed consent (Philippakis *et al.* 2015). Some authors have discussed the tension between privacy and research regarding genomic data around the concept of right to science (Knoppers and Thorogood 2017). In order to facilitate the use of molecular data in research, recommendations have occurred to create digital twins of individuals to computationally test new drugs in the era of precision medicine (Björnsson *et al.* 2019). One solution not stressing privacy might limitedly involve the posthumous donation of medical data on the condition that any potential harm to family members will be monitored (Krutzinna *et al.*, 2019). Proper education of both patients and researchers can contribute to a more general solution. At this point, we should discuss the legal term of informed consent.

Informed consent has been well discussed as a concept, especially in medical law literature. However, its discussions are mostly focused on conflicts regarding malpractice. In this context, the nature of informed consent around health records and genetic data should be different here than the conventional informed consent of medical law because these data can be used in future research, or new findings may require feedback from the participant. Data use is a dynamic process: As the gathered data accumulates, new insights can emerge and secondary use can become routine (Hand 2018). This can make iterative contact with the participant necessary. Therefore, having consent to share genomic data be a dynamic and participatory process can be suggested rather than as a one-time incidence. Of course, continuous participation cannot be compulsory; a participant may choose to give permission to having their genetic data shared at first and then later in the future prefer to not know potential health risks (Berkman and Hull 2014) or not allow further access to their data under any possible scenario. Thus, pre-informed consent should also cover these rights as well. However, foreseeing all possible future uses of data is practically impossible, which is why some authors have suggested that conventional data ownership is not applicable in health related big data and there is a need for a paradigm shift from data ownership to data custodianship (Andanda 2019). Data custodianship is a legal responsibility of data controller, to keep the data safe and maintain access to the data in a fair manner. Therefore in this proposed paradigm, data controller do not own the data but play a role as custodian to provide privacy of the data and to make decisions on future access requests while the data subject still have some rights on the data and delegates the decision-making authority on data access to the custodian. Some authors have even proposed that this dynamism be extended to any personal data on a

commercial cloud and that these data should be considered as a commodity owned by the subject not the data collector (Bouchagiar and Canellopoulou-Bottis 2018). The European Union General Data Protection Regulation (EU GDPR 2016) also explicitly assigns natural persons as data owners. In short, data subject should stay as the data owner but data controller needs more authority for practical reasons. It is known that having innovative investments in health be proportional to patent strength has also been discussed (Williams 2016). Therefore, expecting private investors to go into this field without any pragmatic solution to the issue of data ownership is difficult. From the individual perspective, aside from legally assigning data ownership to each relevant person being hard to do, it doesn't solve all problems. For instance, how well a rural elderly lay-person with a low education level can comprehend all the consequences of consenting to share data is unclear (Uçar and İlkılıç 2019). Thus, public education on data awareness is a must in order to have informed consent really be informed, especially among culturally and educationally diverse populations.

Data awareness is not only an issue for ordinary laymen but also for health professionals. Uncertainty is an important epistemological aspect of big data. Research based on health-related big data has tended to focus more on unveiling the underlying correlations between biological events and entities rather than the causalities between them (Uçar and İlkılıç 2019). Clinicians must also grasp this epistemological discrepancy well, otherwise they may misdirect patients about the stochastic health risks produced by digital health data, which may then lead to misinformed consent instead of informed consent.

Another complication around informed consent regarding the use of genomic data is the potential harm toward family members (PCSBI 2012). Genomic data can inform about risks for certain genetic diseases for relatives. That means consenting to allow the use of genomic data can harm others in terms of discrimination or insurance coverage limitations. On this point, the claim can be made that one's genomic data is not only the individual's and constituent of one's identity but also carries information about close relatives' identities. Thus, the procedure of informed consent for sharing genomic data should be even more complicated.

In the light of all these discussions, what we can propose is to build a participatory digital platform for dynamically managing the process of informed consent. Training all parties (i.e. patients, physicians, researchers, data holders) is crucial but likely not enough. What should be discussed is the use of technical trusted intermediary agencies as the deliverers of consent. Informed consent is strictly personal, yet technical complications require handing this process over to an expert deputy trustee using a digital platform.

Sharing data access with third parties

Once the problems of accessing medical data are solved, attention should be paid to transparency (Kaplan 2020) and sharing fair data access with third parties (Andanda 2019). Transparency is crucial in policies, algorithms, and other tools data collectors and data users use for processing data. This point is not only

important for consent from the data source but also for the reliability of the outcomes from the secondary use of these data. Data science has an interdisciplinary nature, and ill-adaptations can cause fatal results. One of the early genomic signature studies on discovering genomic markers for chemotherapeutic agent selection (Potti *et al.* 2006) was found to have fatal statistical errors (Baggerly and Statistics 2009), and eventually had to be retracted by its authors. For this reason, reproducibility is a key issue to solve in bioinformatics and health informatics, and the most obvious way to improve reproducibility is to improve transparency.

Within the last decade, the amount of research on health-related big data has grown exponentially, and the most common considerations discussed by responsible ethics review committees (ERC) have been privacy and confidentiality, informed consent, fairness and justice, trust, data ownership, and transparency (Ienca *et al.* 2018). The dynamic and technical nature of the ethical problems in digital health, bioinformatics, and biotechnology has complicated ERCs' identification of ethical issues. Having ERC members receive specific trainings and allowing room for experts from the field of data science field in ERCs (or more radically having more specialized data security committees) have been recommended in order to be able to overcome the obstacles (Ienca *et al.* 2018). However, care should be made to not add bricks to the bureaucratic walls. ERCs can easily became a barrier in front of scientific research (Nicholl 2000). In order to find an optimal solution in this trade-off, normative and perhaps somewhat casuistic principles should be adhered to.

Another ethical issue involves the fair provision of access to medical data without any commercial or social discrimination. Most data clouds reside in the northern hemisphere, which means any political conflict with superior world powers may result in an embargo. Important questions to address are the ethicality of limiting access to open medical data from countries under economic sanctions and the ability of an individual to discriminatively limit access to their medical data. Data and access to data are elements of contemporary political power. Within the scope of a political conflict, that power can be used against those who are weaker. We suggest that public good should be interpreted as the good of all humanity and any consent to access medical or genomic data should not be constrained by personal or national political conflicts.

Aside from dystopic governmental scenarios, the monopolization of digital healthcare service providers may cause another ethical problem. As cloud providers or AI-based tool providers create a global oligopoly, the trade-off between commercial interests and public interests may be biased. The claim can be made that a similar case already exists for drug companies. Physicians play a buffer role in favour of individual interests in the conventional scenario, whereas the physician's role has been passivized in the digital era. The oligopoly of data collectors and digital service providers has become dominant. One possible solution for this concern may be distributed ledger systems, as these would break any possible monopolization, at least in the data storage landscape. Like block-chain technologies, these decentralized systems can enhance privacy and contribute to minimizing shared data without limiting the public good. However, data storage constraints will limit their use. Distributed ledger

systems are not suitable for big data; however, extracting a manageable summary (i.e. metadata) of the raw data can be used for this purpose. The use of this technology in bioinformatics is still in its infancy and very promising (Chen and Shae 2019). Federated learning has been proposed for the data analytics portion of AI applications in health informatics to enable algorithm training without needing to share data (Rieke *et al.* 2020). Probable improvements in these kinds of decentralized approaches will help boost data usage without increasing ethical concerns. In addition, updating medical school curricula or providing the appropriate extra trainings to provide additional awareness to physicians may break the passivation process.

Genetic data from biotechnology and synthetic biology

Data-centric digital life sciences have enabled biologists to develop techniques to stretch the frontiers of humankind with new ethical challenges. Inserting or deleting certain genes from organisms using recombinant DNA technologies is routine in life science research but never applied to humans because of ethical concerns. The U.S. Supreme Court's 2013 decision on Myriad Genetics ruled against the ability to patent human-based DNA while ruling for the ability to patent cDNA (synthetic products of complementary DNA molecules; Burk 2014). As a recently developed technique, CRISPR has allowed DNA to be tailored in much higher resolution (Doudna and Charpentier 2014) and brought its developers the 2020 Nobel Prize in Chemistry. However, CRISPR's application on humans was also forbidden due to ethical issues as no certainty exists regarding all the consequences. Nevertheless, one researcher conducted that technique without any ethical permission, and the first genetically modified babies were born a few years ago in China (Cyranoski 2019).

As mentioned in the introduction, J. Craig Venter's claim of being able to create life without any need for a god has definitely created a storm of bioethical debates. The United States' PCSBI discussed this issue and cautiously stated that Venter's accomplishment, 'although extraordinary in many ways, does not amount to creating life as either a scientific or a moral matter' (PCSBI 2010, p. 3).

Apparently, human curiosity will constantly push beyond all ethical barriers. We can foresee that babies will be born who have artificially generated/modified genetic sequences, as well as a flood of other ethical and moral discussions regarding intellectual property rights over genetic sequences to again become a hot topic. When considering another futuristic case where a company develops a genetic modification cascade to obtain genetically enhanced babies, will this genetic data be treated as a commercial good? Upcoming developments will bring genetic data owned not by any individual but by commercial firms. Agriculture is already experiencing these issues on genetically modified organisms. Huge commercial companies already exist in crop sciences for patenting seeds' genetic data. If humanity is to keep considering human genetic data as a part of human heritage (UNESCO 1997), then the commercialization of genetic engineering for humans should not be allowed.

Biomedical devices and data

Rather than being a one-time static issue, wearable biomedical devices have dynamically transformed the issue of the sharing of personal health data, and informed consent standards and data access policies should be regulated accordingly for biomedical devices. Some unique ethical challenges exist around digital medicine, similar to the misconceptions in informed consent forms taken by physicians and user agreements with device producers (Klugman *et al.* 2018). Technical and clinical validation procedures for those processes have also created unique challenges (Mathews *et al.* 2019).

Using data in medical care through clinical decision support software and digital medicine devices creates additional complications for physicians in validating the efficiency of treatments compared to traditional pharmaceutical treatments. Physicians can act as an independent agent in filtering out the capitalistic decisions coming from drug companies to assess a balance. However, in this context of digital health, they need to get additional formations to maintain that role. The dominance of technology provider companies in the field may reduce physicians' and patients' roles, which will definitely create an ethical challenge. What should be recommended is training medical staff accordingly to be aware of advantages or disadvantages of the products and assigning independent expert boards to control the quality of the services. The curricula of medical faculties can also be reviewed from this perspective.

Genetic surveillance

One important ethical challenge in the context of digital life data emerges from the tension between human rights and state power. USA's PCSBI (2012) recommends giving privileges to police and defense forces to access bio-specimens without consent only in exceptional circumstances. We have already encountered debates regarding these exceptional cases, with one USA-based genetic company having shared individuals' genetic data with FBI without informing its customers (Haag 2019). USA governmental organizations have a huge DNA database collected from suspects and convicts; as a debate exists on the tendency of police forces to stop and detain people with a racial bias, the suggestion can be made that these databases will be dominated by racially disadvantaged individuals and that this genetic surveillance mechanism will create further bias against them (Roberts 2011). On the other hand, the European Court of Human Rights convicted the UK in 2008 for not destroying the genetic material of detainees who had ultimately been vindicated (ECHR 2008). We can recommend that, even if individuals are convicts, some limitations must exist for the type of offence and a time limit should also be present to keep the genetic data in parallel with the right to be forgotten.

AI-related ethical issues

Accountability and data reliability are the key points in ethically evaluating AI-based systems. No satisfactory ethical resolution to AI or big data-related

ethical issues can occur if data analytics are allowed to obscure human decisions (Markham *et al.* 2018). All that can be done is to list the ethical problems over and over. However, avoiding responsibility and obscuring the role human decision-making has in the process cannot happen when one highlights the distinction between machines and humans. AI technologies are vulnerable to the inherited bias introduced to them by means of training data sets, and this creates additional concerns about discriminations with the increased use of AI in medical decision making (Char *et al.*, 2018; Schönberger 2019). What should be highlighted on this point is that AI is just a tool like a stethoscope: One cannot properly hear the sounds of the body without it, and it doesn't make any medical decisions independent of the physician. Analogically, AI-based health technologies help physicians capture the details and features they might miss without these tools, but again the medical decision and accountability should stay solely with the physician. AI does not replace or passivize the physician, instead it requires physicians to possess additional configurations akin to using a stethoscope. The human factor in this process has also been highlighted in EU GDPR (2016, Schönberger 2019). Once again, having clinical decision makers be properly trained is crucial to overcome possible novel ethical complications (Rigby 2019). Data reliability is a key concept in improving the efficiency of AI solutions in a fair way. When considering a hypothetical instance such as training an AI algorithm for deciding whether a pregnancy should be finalized with a caesarean section or normal delivery using past data, which is already ethically questionable, this will definitely consolidate unnecessary caesarean section practices. Therefore, ethical principles should be taken into consideration not only in the development process of AI algorithms but also in the training data pre-processing period.

Conclusions

The extent to which our biomolecular data is a part of our identity is a question worth addressing but would exceed the scope of this paper. In this data-centric digital era, not only people's biological status but also their medical condition appear as data, with all personal, social, and financial activity able to turn into digital data. Then 'it is clearly time to reframe what constitutes the boundaries of being' (Markham *et al.* 2018, p. 2), and these frames will shape the ethical tension between individual and public interests. This study is able to summarize its findings and suggestions into six groups. (i) The first problem to address is how the mechanism of informed consent will be able to be efficient in this paradigm. Aside from the need for proper training of all parties, the need exists for expertise as a deputy trustee between individuals and the agencies demanding data in intermediating the dynamic data access permission processes. This could solve the tension between privacy concerns and the need for efficient utilization of data. (ii) Proper training is crucial for data collectors, data holders, physicians, and ethical commissions. Collaborations among the proper experts and re-evaluations of related curricula appear to be vital. (iii) Distributed ledger systems that limits data share can provide privacy to individuals and

prevent unethical commercial and political agendas from being included. (iv) As human curiosity pushes its limits, synthetic biology and biotechnology will tend to invade medical field. We have to stick to UNESCO's declaration that human genetic data is a human heritage and shall not be evaluated as an ordinary good. (v) As EU GDPR (2016) also stated, the human factor must be enhanced in AI systems despite the perceptions of the digital age and the required skills need to be provided to the related experts, especially physicians. Physicians should not be deceived by common misperceptions; instead, they have to be proactive in using digital technologies in medical care services. Physicians must keep playing a buffer role between patients and the companies that provide biomedical technologies or any digital health service. (vi) Even after overcoming privacy and trust problems, data should be clean, neat, reliable, and unbiased toward any discrimination in order to be able to utilize health data efficiently. This problem should be seriously considered by corresponding private and state administrations; otherwise, the fruits of the digital age will remain a fantasy.

Ned Ludd, a weaver afraid of losing his job in the eighteenth century, resisted the introduction of steam-powered engines. Under his leadership of a gang of vandals, the Luddites destroyed many steam-powered looms. In the end, he and his associates turned out to be wrong and to have overreacted. Now, what humanity must do is address all ethical concerns and be well equipped with the skills required of the digital age without being a contemporary Luddite.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The author(s) reported there is no funding associated with the work featured in this article.

Notes on contributors

Dr. Muhammed Erkan Karabekmez has BS, MSc, and PhD (Bogazici University) in Chemical Engineering and LLB (Istanbul University). He studies on systems biology, computational biology, and bioinformatics besides ethical and legislative issues on genomic data.

ORCID

Muhammed Erkan Karabekmez http://orcid.org/0000-0002-0517-5227

References

Andanda, P., 2019. Towards a paradigm shift in governing data access and related intellectual property rights in Big data and health-related research. *IIC - international review of intellectual property and competition law*, 50 (9), 1052–1081. doi:10.1007/s40319- 019-00873-2.

Auffray, C., Charron, D., and Hood, L., 2010. Predictive, preventive, personalized and participatory medicine: back to the future. *Genome medicine*, doi:10.1186/gm178.

Baggerly, K., and Statistics, K.C., 2009. Deriving chemosensitivity from cell lines: forensic bioinformatics and reproducible research in high-throughput biology. *JSTOR*, 3 (4), 1309–1334.

Berkman, B.E., and Hull, S.C., 2014. The "right not to know" in the genomic era: time to break from tradition? *The American journal of bioethics*, 14 (3), 28–31. doi:10.1080/15265161.2014.880313.

Björnsson, B., *et al.*, 2019. Digital twins to personalize medicine. *Genome medicine*, 12 (1), 10–13. doi:10.1186/s13073- 019-0701-3.

Boenink, M., 2010. Molecular medicine and concepts of disease: the ethical value of a conceptual analysis of emerging biomedical technologies. *Medicine, health care and philosophy*, 13 (1), 11–23. doi:10.1007/s11019- 009-9223-x.

Bouchagiar, G., and Canellopoulou-Bottis, M. 2018. Data in the cloud: whose control and whose consent? Available from: https://papers.ssrn.com/abstract=3167033 [Accessed 28 Mar 2021].

Burk, D.L., 2014. The curious incident of the Supreme Court in Myriad genetics. *Notre dame law review*, 90. Available from: https://heinonline.org/HOL/Page?handle=hein.journals/tndl90&id=529&div=&collection= [Accessed 26 Mar 2021].

Char, D.S., Shah, N.H., and Magnus, D., 2018. Implementing machine learning in health care — addressing ethical challenges. *The new England journal of medicine*, 378 (11), 981–983. doi:10.1056/NEJMp1714229.

Chen, Y.-W., and Shae, Z.-Y., 2019. Blockchain for pre-clinical and clinical platform with big data. In: *Application of Omics, AI and Blockchain in bioinformatics research*, 29–46. Available from: https://en.wikipedia.org/wiki/Mirth_ [Accessed 28 Mar 2021].

Colling, R., *et al.*, 2019. Artificial intelligence in digital pathology: a roadmap to routine use in clinical practice. *The journal of pathology*, 249 (2), 143–150. doi:10.1002/path.5310.

Cyranoski, D., 2019. What's next for CRISPR babies. *New England journal of medicine*, 380 (10), 897–899.

Declaration of Helsinki. 1964. *Ethical principles for medical research involving human subjects*, *WMA*. Available from: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research- involving-human-subjects/.

Doudna, J.A., and Charpentier, E., 2014. The new frontier of genome engineering with CRISPR-Cas9. *Science*, 346 (6213), 1258096. doi:10.1126/science.1258096.

ECHR. 2008. *Case of S. and Marper v. the United Kingdom*.

EC Oviedo Convention. 1997. *Convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine: convention on human rights and biomedicine* *.

EU GDPR. 2016. *General data protection regulation (GDPR) – official legal text*. Available from: https://gdpr-info.eu/ [Accessed 28 Mar 2021].

Gubrium, A.C., Hill, A.L., and Flicker, S., 2014. A situated practice of ethics for participatory visual and digital methods in public health research and practice: a focus on digital storytelling. *American journal of public health*, 104 (9), 1606–1614. doi:10.2105/AJPH.2013.301310.

Haag, M. 2019. FamilyTreeDNA admits to sharing genetic data with F.B.I. – the New York Times. *The New York Times*. Available from: https://www.nytimes.com/2019/02/04/business/family-tree-dna-fbi.html [Accessed 31 Mar 2021].

Hand, D.J., 2018. Aspects of data ethics in a changing world: where are we now? *Big data*, 6 (3), 176–190. doi:10.1089/big.2018.0083.

Hongladarom, S., 2006. Ethics of bioinformatics: a convergence between bioethics and computer ethics. *Asian biotechnology and development review*, 9 (1), 37–44.

Ienca, M., *et al.*, 2018. Considerations for ethics review of big data health research: a scoping review. *PLoS one*, 13 (10), 1–15. doi:10.1371/journal.pone.0204937.

J. Craig Venter Institute, 2008. *Scientists create first synthetic bacterial genome – largest chemically defined structure synthesized in the lab – ScienceDaily*. Available from: https://www.sciencedaily.com/releases/2008/01/080124175924.htm [Accessed 31 Mar 2021].

——.. 2010. *First self-replicating synthetic bacterial cell, JCVI*. Available from: https://www.jcvi.org/sites/default/files/assets/projects/first-self-replicating-synthetic-bacterial-cell/press-release-final.pdf [Accessed 31 Mar 2021].

Kaplan, B., 2020. Seeing through health information technology: the need for transparency in software, algorithms, data privacy, and regulation. *Journal of law and the biosciences*, 105 (1), 176–176. doi:10.1353/mon.2013.0025.

Klugman, C.M., *et al.*, 2018. The ethics of smart pills and self-acting devices: autonomy, truth-telling, and trust at the Dawn of digital medicine. *The American journal of bioethics*, 18 (9), 38–47. doi:10.1080/15265161.2018.1498933.

Knoppers, Bartha Maria and Chadwick, Ruth, 2005. Human genetic research: emerging trends in ethics. *Nature reviews genetics*, 6 (1), 75–79. doi:10.1038/nrg1505.

Knoppers, B.M., and Thorogood, A.M., 2017. Ethics and big data in health. *Current opinion in systems biology*, 4, 53–57. doi:10.1016/j.coisb.2017.07.001.

Krutzinna, J., Taddeo, M., and Floridi, L., 2019. *Enabling posthumous medical data donation: an appeal for the ethical utilisation of personal health data, science and engineering ethics*. Springer Netherlands. doi:10.1007/s11948- 018-0067-8.

Lee, L.M., 2017. Ethics and subsequent use of electronic health record data. *Journal of biomedical informatics*, 71, 143–146. doi:10.1016/j.jbi.2017.05.022.

Long, S., and Alpern, R., 2009. Science for future physicians. *Science*, 324 (1241). Available from: https://sci-hub.do/https://science.sciencemag.org/content/324/5932/1241.summary [Accessed 25 Mar 2021].

MAIEI. 2021. *The state of AI ethics, Montreal AI ethics institute*. Available from: http://marefateadyan.nashriyat.ir/node/150.

Markham, A.N., Tiidenberg, K., and Herman, A., 2018. Ethics as methods: doing ethics in the era of big data research—introduction. *Social media and society*, 4, 3. doi:10.1177/2056305118784502.

Martensen, R., 2001. The history of bioethics: an essay review. *Journal of the history of medicine and allied sciences*, 56 (2), 168–175. doi:10.1093/jhmas/56.4.400.

Mathews, S.C., *et al.*, 2019. Digital health: a path to validation. *Npj digital medicine*, 2 (1), 1–9. doi:10.1038/s41746- 019-0111-3.

McGuire, A.L., Caulfield, T., and Cho, M.K., 2008. Research ethics and the challenge of whole-genome sequencing. *Nature reviews genetics*, 9 (2), 152–156. doi:10.1038/nrg2302.

Moor, J.H., 1985. What is computer ethics? *Metaphilosophy*, 16 (4), 266–275.

Myers, J., *et al.*, 2008. Privacy and public health at risk: public health confidentiality in the digital age. *American journal of public health*, 98 (5), 793–801. doi:10.2105/AJPH.2006.107706.

Nebeker, C., *et al.*, 2017. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. *AJOB empirical bioethics*, 8 (4), 266–276. doi:10.1080/23294515.2017.1403980.

Nicholl, J., 2000. The ethics of research ethics committees. *BMJ (clinical research ed.)*, 320 (7243), 1217. doi:10.1136/bmj.320.7243.1217.

Panch, T., Szolovits, P., and Atun, R., 2018. Artificial intelligence, machine learning and health systems. *Journal of global health*, 8 (2), 1–8. doi:10.7189/jogh.08.020303.

PCSBI. 2010. *New directions: the ethics of synthetic biology and emerging technologies, Presidential Commission for the Study of Bioethical Issue*. Available from: www.bioethics.gov.

——.. 2012. *Privacy and progress in whole genome sequencing, Presidential Commission for the Study of Bioethical Issues*. doi:10.1515/jwiet- 2014-0119.

Pesapane, F., *et al.*, 2018. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. *Insights into imaging*, 9 (5), 745–753. doi:10.1007/s13244- 018-0645-y.

Peters, T., 2018. Playing god with Frankenstein. *Theology and science*, 16, 145–150. doi:10.1080/14746700.2018.1455264.

Philippakis, A.A., *et al.*, 2015. The matchmaker exchange: a platform for rare disease gene discovery. *Human mutation*, 36 (10), 915–921. doi:10.1002/humu.22858.

Potti, A., *et al.*, 2006. Genomic signatures to guide the use of chemotherapeutics. *Nature medicine*, 12, 1294–1300. Available from: https://www.nature.com/articles/nm1491 [Accessed 28 Mar 2021].

Rieke, N., *et al.*, 2020. The future of digital health with federated learning. *Npj digital medicine*, 3 (1), 1–7. doi:10.1038/s41746- 020-00323-1.

Rigby, M.J., 2019. Ethical dimensions of using artificial intelligence in health care. *AMA journal of ethics*, 21 (2), E121–E124. doi:10.1001/amajethics.2019.121.

Rippen, H., and Risk, A., 2000. e-Health code of ethics (May 24). *Journal of medical internet research*, 2 (2), 1–6. doi:10.2196/jmir.2.2.e9.

Roberts, D., 2011. Collateral consequences, genetic surveillance, and the new biopolitics of race. *Howard law journal*, 54, 12–54.

Ruotsalainen, P., and Blobel, B., 2020. Health information systems in the digital health ecosystem—problems and solutions for ethics, trust and privacy. *International journal of environmental research and public health*, 17 (9), 1–15. doi:10.3390/ijerph17093006.

Schönberger, D., 2019. Artificial intelligence in healthcare: a critical analysis of the legal and ethical implications. *International journal of law and information technology*, doi:10.1093/ijlit/eaz004.

Shah, P., *et al.*, 2019. Artificial intelligence and machine learning in clinical development: a translational perspective. *Npj digital medicine*, 2 (1), doi:10.1038/s41746- 019-0148-3.

Shameer, K., *et al.*, 2017. Translational bioinformatics in the era of real-time biomedical, health care and wellness data streams. *Briefings in bioinformatics*, 18 (1), 105–124. doi:10.1093/bib/bbv118.

TCGA. 2021. Available from: https://portal.gdc.cancer.gov/ [Accessed 3 Apr 2021].

Tennison, M.N., and Moreno, J.D., 2012. Neuroscience, ethics, and national security: the state of the art. *PLoS biology*, 10 (3), 3–6. doi:10.1371/journal.pbio.1001289.

Uçar, A., and İlkılıç, İ, 2019. Büyük Verinin Sağlık Hizmetlerinde Kullanımında Epistemolojik ve Etik Sorunlar. *Istanbul University institute of health sciences journal of advanced research in health sciences*, 2 (2), 80–92. doi:10.26650/jarhs2019- 616389.

UNESCO. 1997. *Universal declaration on the human genome and human rights*. Available from: http://portal.unesco.org/ [Accessed 3 Apr 2021].

——. 2003. *International declaration on human genetic data*. Available from: http://portal.unesco.org/ [Accessed 3 Apr 2021].

van den Belt, H., 2009. Playing god in Frankenstein's footsteps: synthetic biology and the meaning of life. *Nanoethics*, 3 (3), 257–268. doi:10.1007/s11569- 009-0079-6.

Venter, J.C., 2014. *Life at the speed of light: from the double helix to the dawn of digital life*. New York: Penguin.

Williams, H.L., 2016. Intellectual property rights and innovation: evidence from health care markets. *Innovation policy and the economy*, doi:10.1086/684986.

Wjst, M., 2010. Caught you: threats to confidentiality due to the public release of large-scale genetic data sets. *BMC medical ethics*, 11 (1), 2–5. doi:10.1186/1472- 6939-11-21.