sciendo

Artificial intelligence in healthcare: Threats to the fundamental values of our society

Klára Zikmundová*

Summary: The article deals with the ethical and legal aspects of the use of artificial intelligence in healthcare. The author analyses key threats to the fundamental values held by the majority of democratic society and the human rights guaranteed by international human rights treaties, such as the European Convention on Human Rights. Emphasis is placed, in particular, on the risk of restricting patients' autonomy due to an insufficient regulation of the use of clinical decision-support systems, a lack of humanistic care due to excessive use of care robots, a decline in human skills in specific fields of medicine, such as radiology, and the violation of privacy by extensive surveillance of social media. The article provides an overview of the benefits of artificial intelligence for healthcare, but also warns of the downsides, with the aim to encourage a debate that would lead to reasonable legal regulation.

Keywords: Artificial intelligence, human rights, patient autonomy, privacy, surveillance, healthcare.

Citation: ZIKMUNDOVÁ, K. Artificial intelligence in healthcare: Threats to the fundamental values of our society. *European Studies – the Review of European law, Economics and Politics*. 2022, vol. 9, no. 2, pp. 181–196 DOI: 10.2478/eustu-2022-0019.

Ackowledgements: This article was written within project no. CZ.02.2.69/0.0/0.0/19_073/0016935: START/SOC/017, titled Law and Artificial Intelligence – Challenges and Opportunities, implemented at the Faculty of Law, Charles University.

* Klára Zikmundová is a PhD candidate at Faculty of Law, Charles University, Prague, Czech Republic, e-mail: zikmunkl@prf.cuni.cz.

1. Introduction

Artificial Intelligence (AI), as a *„computer system that can do tasks that humans need intelligence to do,"*[1] is already used in the delivery of healthcare and can be expected to expand even further into many fields of healthcare in the future.

A variety of AI-powered medical tools, such as diagnosing tools in the field of radiology and treatment recommendation systems, are being developed. Many experts see the emerging technologies as a huge benefit to humanity, particularly in medicine, as AI-powered tools are likely to have a significant impact on the satisfaction of everyone's right to the highest attainable standard of health.[2]

Other experts, however, point out the threats of AI. The foreseeable negative effects of AI range, in general, from social issues, such as unemployment, to threats to human rights, such as the deprivation of privacy,[3] increased bias and the ensuing inequality, and dominance of AI over humans, ultimately leading to the extinction of humanity.[4] Nevertheless, given the latest advancements in medical AI, the latter problem appears to be quite distant.

This article focuses on AI in healthcare and its potential impacts, including ethical considerations. Do AI-powered medical tools pose a threat to the fundamental values of our society? How can we contribute to the development of AI today, so that it shapes the world we want to live in tomorrow?

The article aims to provide an overview of the key threats to our values ensuing from the use of AI in healthcare.

2. Background

AI-powered tools might be highly beneficial in many areas of healthcare. For instance, AI can be used for electronic health record systems (e.g. Electronic Data Capture system),[5] diagnostics and treatment recommendations (e.g. Watson for Oncology),[6]

---

1 O'NEILL, M. *Explainer: what is artificial intelligence?* [online]. Available at: http://www.abc.net.au/news/2017-08-07/explainer-what-is-artificial-intelligence/8771632

2 RASO, F. A., HILLIGOS, H. et al. *Artificial Intelligence & Human Rights: Opportunities & Risks*. Berkman Klein Center for Internet & Society at Harvard University, Research Publication no. 2018–6, 2018.

3 Ibid.

4 BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014.

5 MONTE, R. *The Beginner's Guide to an Electronic Data Capture (EDC) System* [online]. Available at: https://www.advarra.com/blog/beginners-guide-electronic-data-capture-edc-system/

6 SOMASHEKHAR, S. P. et al. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. *Journal of Clinical Oncology*. 2017, vol. 35, pp. 8527–8527.

care robots[7] and surgical robots.[8] There is a variety of technologies applied, ranging from knowledge-based systems to deep neural networks. Knowledge-based systems are able to make recommendations through a series of encoded conditional statements. These systems already appeared in 1970s–1980s, with MYCIN – a system recognising bacterial infection and suggesting treatment – as the most well-known specimen.[9]

Nevertheless, knowledge-based systems, which were unable to solve complex problems, were *„largely abandoned for a new approach to AI called machine learning, which seeks to relate variables to outcomes through complex non-linear functions learned directly from the data (without necessarily having prior domain knowledge)"*[10] in 1990s. Even more sophisticated deep neural network systems are on the rise today.

One of the most famous AI-powered medical tools is IBM's Watson for Oncology, already mentioned above. IBM's intention was to develop a tool that would make treatment recommendations by extracting patient data from medical records and evaluating this data on the basis of an expertise derived from medical journals and textbooks accessible to the system. The result was supposed to be a ranked list of treatment options. However, this attempt failed.[11]

Diagnostic tools seem to be more promising. For instance, AI can be utilised in dermatology to classify suspicious skin lesions and thus detect skin cancer. Some studies have yielded encouraging results and, according to some experts, these tools might even outperform human clinicians.[12]

Regarding diabetic retinopathy, the United States Food and Drug Administration has already approved the JDx-DR device, which can provide a screening decision without even having to be interpreted by a human clinician.[13]

---

7 COGHLAN, S. Robots and the Possibility of Humanistic Care. *International Journal of Social Robotics*. 2022, vol. 14, pp. 2095–2108.

8 AMISHA, P. M., PATHANIA, M., RATHAUR, V. K. Overview of artificial intelligence in medicine. *Journal of family medicine and primary care*. 2019, vol. 8, no. 7, pp. 2328–2331.

9 BENNET, C. C., DOUB, T. Expert Systems in Mental Healthcare: AI Applications in Decision Making and Consultation. In: LUXTON, D. (ed.). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier, 2016, pp. 27–47.

10 QUINN, T. et al. The three ghosts of medical AI: Can the black-box present deliver? *Artificial Intelligence in Medicine*. 2022, vol. 124.

11 *Artificial Intelligence Failure at IBM ,Watson for Oncology'* [online]. Available at: https://icmrindia.org/casestudies/catalogue/IT%20and%20Systems/ITSY126.htm

12 ESTEVA, A., KUPREL, B. et al. Dermatologistlevel classification of skin cancer with deep neural networks. *Nature*. 2017, vol. 7639, no. 542, pp. 115–118.

13 FROOMKIN, A. M., KERR, I., PINEAU, J. When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning. *Arizona Law Review*. 2019, vol. 61, no. 33, pp. 32–99.

A number of benefits are brought by care robots, which are *„designed for use in home, hospital, or other settings to assist in, support, or provide care for sick, disabled, young, elderly, or otherwise vulnerable persons"*.[14] Such robots can provide physical care such as lifting, feeding, cleaning and bathing. On top of that, some experts even encourage development of affective care robots, which are capable of comforting, reassuring, and giving solace.[15]

AI influences many fields of healthcare, and mental health is no exception. For instance, SimSensei Kiosk is a virtual human interviewer designed to detect indications of depression, anxiety or post-traumatic stress disorder (PTSD).[16] According to some experts, such a tool makes end users more comfortable to talk about their mental health issues than when speaking to a human psychologist.[17]

Further, AI could be used for preliminary screening of social media to detect mental health disorders. This would make it possible to deliver timely and appropriate treatment or even predict mass shootings.[18]

Although the above-mentioned AI applications in healthcare might be highly beneficial for humanity, they raise to a number of questions. I shall provide an overview of the major concerns in this article.

3. Patient autonomy

A paternalistic approach to the clinician-patient relationship prevailed in the past. The physician was considered a guardian of the patient's health and decided unilaterally what treatment was in the patient's best interest. This approach, however, has mostly given way to the currently prevailing principle of patient autonomy.[19]

Patient autonomy means, among others, that patients should be allowed to make treatment decisions based on their own values. It goes without saying that

[14] VALLOR, S. Carebots and Caregivers: Sustaining the Ethical Ideal of Care in the Twenty-First Century. *Philosophy & Technology.* 2011, vol. 24, art. no. 251.

[15] COGHLAN, S. Robots and the Possibility of Humanistic Care. Op. cit.

[16] DEVAULT, D. et al. SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support. *13th International Conference on Autonomous Agents and Multiagent Systems*. 2014, pp. 1061–1068.

[17] RIZZO, A. et al. Autonomous Virtual Human Agents for Healthcare Information Support and Clinical Interviewing. In: LUXTON, D. (ed.). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier, 2016, pp. 53–79.

[18] NEUMAN, Y. Artificial Intelligence in Public Health Surveillance and Research. In: LUXTON, D. (ed.). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier, 2016, pp. 231–254.

[19] CHIN, J. J. Doctor-patient relationship: from medical paternalism to enhanced autonomy. *Singapore Medical Journal.* 2002, vol. 43, pp. 152–155.

the patient must be able to make an informed decision. It is therefore crucial that the clinician communicate the treatment options and their advantages and disadvantages to the patient in an understandable way.[20]

An emphasis on patient autonomy has been growing in the Western world, with simultaneous development of AI-powered clinical decision support systems (CDSS). Do such systems pose a threat to patient autonomy?

Generally speaking, CDSS is a system designed to improve healthcare delivery by improving healthcare decisions through the use of targeted expert knowledge and patient clinical data. Some of these systems may help clinicians with diagnosis, while others are developed to recommend treatment options. To achieve the best results, the outcomes should then be combined with human clinician expert knowledge.[21]

There are two types of CDSS: knowledge-based systems and systems based on machine learning and deep neural networks. The latter are particularly promising in terms of healthcare development, but they also raise ethical concerns. *„Non-knowledge based CDSS, although a rapidly growing use case for AI in medicine, are rife with challenges including problems understanding the logic that AI uses to produce recommendations (black boxes), and problems with data availability.“*[22]

McDougall perceives CDSS as a threat to patient autonomy and a potential inclination towards a paternalistic approach in medicine. The treatment options suggested by CDSS do not reflect the values and interests of the individual patient, but rather values that are generally accepted, while the values and preferences of individual patients can vary significantly. For example, while some terminal patients prefer to live as long as possible, others put the quality of life above its length.[23]

Hence, CDSS pose the risk that clinicians will consider the treatment options generated by the system to be the best options available, even though they may not match the individual patients' values.[24] It can be countered that such tools are able to suggest a prioritised list of treatment options, which should be subsequently consulted with the patient; the patient can then make an informed decision on the preferred option.

---

[20] ENTWISTLE, V. A. Supporting Patient Autonomy: The Importance of Clinician-patient Relationships. *Journal of General Internal Medicine.* 2010, vol. 25, no. 7, pp. 741–745.

[21] SUTTON, R. T., PINCOCK, D., BAUMGART, D. C. et al. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digital Medicine*. 2020, vol. 3, art. no. 17.

[22] Ibid.

[23] MCDOUGALL, R. Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*. 2019, vol. 45, no. 3, pp. 156–160.

[24] Ibid.

Nonetheless, it is necessary to promote the principle of patient autonomy in the context of the use of CDSS, and to prevent the clinician from autonomously determining the most appropriate treatment for the patient. Clinicians, of course, have expert knowledge and experience that makes them qualified to make medical decisions. But they have to take into account the values of each patient, which need not necessarily correspond to the values generally accepted by society. With the use of medical AI, more emphasis should be paid to informing patients and incorporating ethical subjects in the curricula of medical students.

Another threat ensuing from the use of AI-powered CDSS is the influence of stakeholders, such as pharmaceutical companies, on treatment. For instance, a CDSS manufacturer could encode hidden values into the algorithm with the aim to encourage the sale of certain medical devices or drugs.[25]

This can be avoided by placing requirements on the transparency and interpretability of algorithms. It could be argued, however, that this is prevented by the complexity of deep learning systems.[26] Such systems, while very powerful, tend to be opaque. This is why they are often referred to as *„black boxes"*.[27]

A certain degree of transparency and explainability is specifically essential in the field of medicine, both to ensure safety and trust, and to enable shared decision-making between the clinician and the patient. Quinn et al. noted: *„We are at a crossroads. Do we continue to extend these powerful but intrinsically opaque models? Or do we prioritize transparency and develop new algorithms that align with the needs of the clinicians and patients who will use them? We (...) argue that opaque models (1) lack quality assurance, (2) fail to elicit trust, and (3) restrict clinician-patient dialogue. We posit that healthcare should mandate transparency in all aspects of model design and model validation in an effort to help ensure the reliability of medical AI."*[28]

I consider it important to have a defined certification process that every CDSS will have to undergo before it can be deployed in healthcare. A generally recognised concern of AI is its opacity, and it would thus be quite problematic to insist on a comprehensive explainability of the algorithm to the end user, i.e. the clinician and the patient. This is because explaining the functionality of CDSS and their rationale, and scenarios in particular, is a technically challenging task.[29]

[25] LIU, X., KEANE, P. A., DENNISTON, A. K. Time to regenerate: the doctor in the age of artificial intelligence. *Journal of the Royal Society of Medicine*. 2018, vol. 111, no. 4, pp. 113–116.

[26] WACHTER, S., MITTELSTADT, B., RUSSELL, Ch. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology*. 2018, vol. 31, no. 2, pp. 841–887.

[27] QUINN, T. et al. The three ghosts of medical AI: Can the black-box present deliver? Op. cit.

[28] Ibid.

[29] WACHTER, S., MITTELSTADT, B., RUSSELL, Ch. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. Op. cit.

That being said, we can insist on comprehensive transparency and explainability of the system's operation to an expert during the certification process.

Explainability should also be ensured for clinicians and patients, but in a different sense. The clinician and the patient need not understand the complex workings of the algorithm.[30] Explainability of a CDSS to clinicians and patients should be seen in terms of the purpose the system is intended to serve, such as enabling a review of the proposed treatment options and shared decision-making.

The certification process is regulated in the European Union by Regulation (EU) 2017/745 on medical devices.[31] The Regulation deals with all medical devices, not only AI. AI-powered CDSS fall within the scope of the regulation given the definition of *„medical device"*. However, the Regulation does not address the challenges of AI explainability.

4. Humanistic care

We can already see robots at work in the healthcare sector. For example, the HelpMate service robot is used in U.S. hospitals to transport various items, from bed linen to medicines.[32]

Surgical robots represent another way of using robots in healthcare. These robots are used in surgery, but do not operate autonomously, as they need a direct input for each task from a human surgeon. For illustration, the DaVinci system makes it possible to perform surgeries and related procedures more accurately and quickly than if they were done by a sole human surgeon, using just one incision.[33]

In the near future, we can expect rise of care robots performing nursing tasks, such as bathing, lifting, and feeding. Wynsberghe suggests that care robots are a feasible solution to the shortage of healthcare personnel, which is expected to become an increasingly severe concern given the population ageing and low fertility rates. She also argues that care robots might mitigate the risk of patient

[30] WACHTER, S., MITTELSTADT, B., RUSSELL, Ch. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. Op. cit.

[31] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017, on medical devices, as amended.

[32] ENGELBERGER, G. HelpMate, a service robot with experience. *Industrial Robot*. 1998, vol. 25, no. 2, pp. 101–104.

[33] DEVLIN, H. *The Robots Helping NHS Surgeons Perform Better, Faster and For Longer*. [online]. Available at: <http://www.theguardian.com/society/2018/jul/04/robots-nhs-surgeons-keyhole-surgery-versius>

abuse or maltreatment, which unfortunately occurs, for example, among the elderly or patients suffering from dementia.[34]

However, there is a number of ethical concerns related to care robots, such as the decrease in the moral quality of care. Decker states: *„People requiring care may not be made into objects by replacing staff with robots. We recommend that robots be employed only as tools or as technical assistants in caregiving and to maintain the autonomy of the care recipient in his/her social environment.“*[35]

Caring for the elderly or sick should comprise not only physical tasks, such as lifting or bathing, but also tasks seeking to promote care-receiver's emotional wellbeing and to reduce the feelings of loneliness and anxiety, as it is commonly recognised that the emotional state has a positive effect on healing.[36] It should be underlined in this context that in Article 12 of the International Covenant on Economic, Social and Cultural Rights (ICESCR), the States Parties recognised the right of everyone to the enjoyment of the highest attainable standard of physical and mental health.

Sparrow, too, doubts the replaceability of human caregivers by care robots, as they lack the empathy, which is necessary for quality care.[37]

Wynsberghe then distinguishes goal-directed and practice-oriented care activities: *„Goal-directed activities are those in which the main point is a state of affairs that is external to the activity. In contrast, practice-oriented activities are those in which the main point is, or heavily depends, on the performance of the activity itself, i.e. the goal is internal to the activity.“*[38]

Caregiving activities usually have two sides. They can be seen as goal-directed and as practice-oriented. However, there are activities where the goal predominates, such as the collection of urine samples.[39] Those activities can then be more likely to be left to care robots.

The deployment of care robots in healthcare seems inevitable. Nevertheless, in my opinion, it is necessary to enact rules stating (i) which care activities are allowed to be performed by care robots; (ii) which care maybe provided by robots only with the patient's specific consent; and (iii) which tasks must remain with human caregivers. Further, it is essential to promote an ethical design in the development of care robots.

---

[34] WYNSBERGHE, A. *Healthcare Robots. Ethics, Design and Implementation*. Routledge: London, 2015, pp. 2–3.

[35] DECKER, M. Caregiving robots and ethical reflection: the perspective of interdisciplinary technology assessment. *AI & Society*. 2008, vol. 22, pp. 315–330.

[36] COGHLAN, S. Robots and the Possibility of Humanistic Care. Op. cit.

[37] SPARROW, R., SPARROW, L. In the hands of machines? The future of aged care. *Minds and Machines*. 2006, vol. 16, no. 2, pp. 141–161.

[38] WYNSBERGHE, A. *Healthcare Robots. Ethics, Design and Implementation.* Op. cit.

[39] WYNSBERGHE, A. When Should We Use Care Robots? The Nature-of-Activities Approach. *Science and Engineering Ethics*. 2016, vol. 22, no. 6, pp. 1745–1760.

5. Human in the loop

Diagnostics are another field where medical AI can be utilised; it turns out that AI has the potential to achieve even better results than humans in this area.

For instance, AI can successfully detect skin cancer. The machine can be trained to recognise different types of skin lesions by uploading a large number of input images (training data) with accompanying information from human clinicians. Subsequently, the system can classify the skin lesions for individual patients. Research carried out in 2017 showed that AI was able to classify skin lesions with the same success rate as experienced dermatologists.[40] In another research done in the same year, AI was even more successful than human dermatologists.[41]

Similarly, radiology is also considered a field particularly suited to the deployment of AI.[42] There are many reasons why this is so. First, analysing CT scans and other images by human radiologists is time consuming and cost ineffective. Second, the assessment of images by AI is more accurate, because AI can spot a connection that a human clinician might not realise. Chockley explains: *„(…) machine learning combines predictors in nonlinear and interactive ways. A human reading a chest radiograph may be inclined to simply interpret the image, determining if it represents someone healthy or sick and, if sick, whether infection, fluid, tumor, or another issue is present. On the other hand, a machine will treat each tiny pixel on the scan as an individual variable and will seek to organize those pixels into shapes and patterns and, from there, make a diagnosis."*[43]

It has been argued that AI will outperform human clinicians,[44] and even that radiologists are about to become redundant.[45]

Medical AI is currently designed to facilitate the work of human clinicians, not to replace them. Also, AI development is extremely costly. However, once an AI tool is developed and deployed in practice, the cost of providing a service to patients is reduced compared to a service provided by a human clinician.

---

[40] ESTEVA, A., KUPREL, B. et al. Dermatologistlevel classification of skin cancer with deep neural networks. Op. cit.

[41] MUKHERJEE, S. *A.I. Versus M.D. What happens when diagnosis is automated?* [online]. Available at: <https://www.newyorker.com/magazine/2017/04/03/ai-versus-md>

[42] FROOMKIN, A. M., KERR, I., PINEAU, J. When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning. Op. cit.

[43] CHOCKLEY, K., EZEKIEL, E. The End of Radiology? Three Threats to the Future Practice of Radiology. *Journal of the American College of Radiology*, 2016, vol. 13, pp. 1415–1420.

[44] RASO, F. A., HILLIGOS, H. et. al. *Artificial Intelligence & Human Rights: Opportunities & Risks*. Op. cit.

[45] MUKHERJEE, S. *A. I. Versus M.D. What happens when diagnosis is automated?* [online]. Available at: https://www.newyorker.com/magazine/2017/04/03/ai-versus-md

Froomkin warns that healthcare providers might reduce the staff of human clinicians in certain fields, especially radiology. This could lead to negative consequences, such as a loss of human skills and expert knowledge.[46]

These are crucial to the development of future generations of medical technology and to identifying the shortcomings of the current ones. It is difficult to predict whether AI will have an impact on the number of human clinicians and the reduction in skills, or whether the opposite will be the case. But this option also has to borne in mind.

6. Respect for privacy

Emergent technologies can also be leveraged for surveillance, either for good or bad purposes. For instance, Neuman suggests AI could be used for preliminary screening for mental health disorders in society. He argues that software based on natural language processing could detect depression from a person's texts on social media. Similarly, other mental health disorders, such as Parkinson disease, Alzheimer's disease, or PTSD, could be detected early and appropriate treatment delivered. Neuman also mentions the possibility of predicting mass shootings based on social media screening.[47]

When it comes to public mental health surveillance and profiling, a distinction needs to be made between surveillance of public speech on social media and monitoring private communications.

The right to respect for private and family life guaranteed by the European Convention on Human Rights (ECHR) would be violated by monitoring private communications. Nevertheless, the ECHR permits interference with this right, *inter alia* on grounds of national security, public safety or crime prevention.[48]

Hence, this chapter will be divided into two parts. The first part will focus on surveillance conducted to prevent crime (e.g. mass shootings) and the second part will deal with surveillance aimed at early detection of mental illnesses and providing advice.

6.1. Crime prevention

The European Court of Human Rights already recognised in 1978 that secret surveillance of private communications is permissible in certain circumstances.

[46] FROOMKIN, A. M., KERR, I., PINEAU, J. When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning. Op. cit.

[47] NEUMAN, Y. Artificial Intelligence in Public Health Surveillance and Research. Op. cit.

[48] European Convention on Human Rights, Article 8.

The court stated: *„The Court has therefore to accept that the existence of some legislation granting powers of secret surveillance over the mail, post and telecommunications is, under exceptional conditions, necessary in a democratic society in the interests of national security and/or for the prevention of disorder or crime."*[49]

The court naturally did not comment on the issue of monitoring communications made via social media, as no such thing existed at the time. However, one cannot see why the above reasoning should not apply to social media. Given the admissibility of monitoring private communications, it is all the more necessary to allow surveillance of public speech on social networks. This applies to security and crime prevention; it is not possible to monitor private communications for purposes other than those set out in Article 8 of the ECHR, for instance, for the general improvement of public mental health.

Similar conclusions regarding the lawfulness of the surveillance of private communications were reiterated by the European Court of Human Rights in the 2015 Roman Zakharov v. Russia and 2016 Cevat Özel v. Turkey decisions. However, the Court stressed that when monitoring private communications, the person concerned had to be given sufficient safeguards to protect their rights.[50] [51]

Compared to historical technologies used to monitor communications, AI enables technically much more extensive and sophisticated surveillance. This is also reflected in the proposal of the Artificial Intelligence Act,[52] which includes a ban on the use of real-time remote biometric identification systems in publicly accessible spaces.

In my view, the permissibility of extensive AI surveillance should be discussed and assessed at a supranational level, in order to provide sufficient guarantees for respecting private correspondence.

6.2. Public mental health

Neuman suggests that AI (e.g. natural language processing) could be used to improve public mental health by screening social media and offering appropriate advice and care to social media users. Although the idea seems to be well-intentioned, it poses a number of threats, especially to the right to privacy under Article 8 of the ECHR.

[49] Klass and Others v. Germany, European Court of Human Rights, 1978.

[50] Roman Zakharov v. Russia, European Court of Human Rights, 2015.

[51] Cevat Özel v. Turkey, European Court of Human Rights, 2016.

[52] Proposal for a regulation of the European Parliament and of the council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts (COM/2021/206).

The boundaries between private and public life are gradually blurred as contemporary society tends to deeply rely on social media. It is not surprising that many companies use AI to predict one's interests, preferred products, or even electoral decisions based on social media profiles.[53]

Protection of data subjects' rights should therefore not be underestimated. There are various cases when publicly accessible social media data materially affected one's life. A good example is the case of *„a Facebook security analyst who had ‚Facebook stalked' a young woman he had met on Tinder and told her about it in a Tinder private message. His objective was to comment on how she was hard to find. The young woman took a screenshot of the private message and tweeted it publicly to Facebook. Facebook found the public message and immediately fired the individual for abusing access privileges."*[54]

Processing of publicly accessible social media data must comply with Regulation (EU) 2016/679 (GDPR), regardless of whether or not the processor processes the personal data for business purposes.[55] It should also be noted that although not expressly stated, a public mental health surveillance tool would probably be used by private businesses in the mental health sector to promote the mental healthcare provider's services.

First, under Article 14 of the GDPR, in cases where personal data were not obtained from the data subject, the processor must provide the subject with transparency information. This would also apply to collecting publicly accessible data from social media.

Second, publishing personal data in social media cannot be understood as consent of the data subject to third-party processing and to any other purpose of processing than promoting one's profile. This was already expressed in 2011 by the French Data Protection Authority (the „DPA").[56]

However, the DPA's decision does not apply the GDPR as the latter did not exist yet at the time. It is based on the Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data[57] and on French law.

In the given case, a French company provided a service showing personal data collected on various sites (e.g. Facebook, Viadeo, Linkedin, Twitter, etc.) on

[53] *Effectively Utilizing Publicly Accessible Social Media Data While Staying GDPR Compliant.* [online]. Available at: https://www.finextra.com/blogposting/16206/effectively-utilizing-publicly-accessible-social-media-data-while-staying-gdpr-compliant

[54] Ibid.

[55] Regulation (EU) 2016/679 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), as amended.

[56] Resolution 2011–203, Commission Nationale de l'Informatique et des Libertés [online]. Available at: https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000024583206/

[57] Council of Europe treaty, 1981.

the basis of a search for a single surname. The company argued that any person who voluntarily provided information about him/herself on social media profile and did not use tools to restrict access to his/her profile could not be unaware that this information might be displayed on the internet, in particular on the results pages of search engines.

The DPA stated that the mere fact that personal data were publicly posted on social media did not allow a third party to proceed with massive, repetitive and undifferentiated collection of these data without informing the persons concerned.

Further, the DPA added that a person's intention to appear on the web and to be indexed by search engines, inferred from the fact that certain data was in a social media profile, *„cannot obviously be imputed to minors. Indeed, it is known that the younger generations are rarely aware of the scope of the dissemination of personal information by social networks, which they are also the most fervent users. To attribute to them a positive intention that their information be included as widely as possible on the network, including on an enhanced directory service, seems excessive in this context. On the contrary, such a collection of personal data is all the more problematic as it concerns a young and vulnerable population, which it is essential to protect."*[58]

I tend to agree with the view expressed by the DPA. Mental health screening should only be allowed with consent of the person concerned, even though such surveillance focuses on publicly available information. The subject has made the information public for the purpose of keeping his/her friends informed about his/her private life and possibly to gain some followers, and not to receive healthcare services.

7. Conclusion

Artificial intelligence applications range from tools designed to simplify administrative processes in hospitals to diagnostic tools that outperform human clinicians. These emergent technologies offer solutions to a number of ongoing challenges, such as the lack of medical staff. They even provide more accurate and faster diagnostics than if the latter were performed by human clinicians.

On the other hand, they pose a threat to our values. It is necessary to ensure that their use does not compromise shared decision-making between the patient and the clinician, and result in a move towards a paternalistic approach

[58] Resolution 2011–203, Commission Nationale de l'Informatique et des Libertés. [online]. Available at: https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000024583206/

to medicine. Furthermore, we can never allow the loss of human expertise and skills in certain fields of medicine, especially radiology, and we need to insist on a respect for privacy, even in the public space created by social media. Last but not least, we should promote humanistic care, i.e. care provided with empathy and understanding.

Despite all the above, we support the development and deployment of AI in healthcare. In this article, we wanted to highlight some of the downsides and contribute to a debate that would lead to reasonable legal regulation.

List of references

AMISHA, P. M., PATHANIA, M., RATHAUR, V. K. Overview of artificial intelligence in medicine. *Journal of family medicine and primary care*. 2019, vol. 8, no. 7, pp. 2328–2331.

BENNET, C. C., DOUB, T. Expert Systems in Mental Healthcare: AI Applications in Decision Making and Consultation. In: LUXTON, David (ed.). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier, 2016, pp. 27–47.

BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, 2014, 328 p.

COGHLAN, S. Robots and the Possibility of Humanistic Care. *International Journal of Social Robotics*. 2022, vol. 14, pp. 2095–2108.

DECKER, M. Caregiving robots and ethical reflection: the perspective of interdisciplinary technology assessment. *AI & Society*. 2008, vol. 22, pp. 315–330.

DEVAULT, D. et al. SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support. *13th International Conference on Autonomous Agents and Multiagent Systems*. 2014, pp. 1061–1068.

ENGELBERGER, G. HelpMate, a service robot with experience. *Industrial Robot*. 1998, vol. 25, no. 2, pp. 101–104.

ENTWISTLE, V. A. Supporting Patient Autonomy: The Importance of Clinician-patient Relationships. *Journal of General Internal Medicine.* 2010, vol. 25, no. 7, pp. 741–745.

ESTEVA, A., KUPREL, B. et al. Dermatologistlevel classification of skin cancer with deep neural networks. *Nature*. 2017, vol. 7639, no. 542, pp. 115–118.

FROOMKIN, A. M., KERR, I., PINEAU, J. When AIs Outperform Doctors: Confronting the Challenges of a Tort-Induced Over-Reliance on Machine Learning. *Arizona Law Review*. 2019, vol. 61, no. 33, pp. 32–99.

CHIN, J. J. Doctor-patient relationship: from medical paternalism to enhanced autonomy. *Singapore Medical Journal,* 2002, vol. 43. pp. 152–155.

CHOCKLEY, K., EZEKIEL, E. The End of Radiology? Three Threats to the Future Practice of Radiology. *Journal of the American College of Radiology*. 2016, vol. 13, pp. 1415–1420.

LIU, X., KEANE, P. A., DENNISTON, A. K. Time to regenerate: the doctor in the age of artificial intelligence. *Journal of the Royal Society of Medicine*. 2018, vol. 111, no. 4, pp. 113–116.

MCDOUGALL, R. Computer knows best? The need for value-flexibility in medical AI. *Journal of Medical Ethics*. 2019, vol. 45, no. 3, pp. 156–160.
NEUMAN, Y. Artificial Intelligence in Public Health Surveillance and Research. In: LUXTON, D. (ed.). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier, 2016, pp. 231–254.
QUINN, T. et al. The three ghosts of medical AI: Can the black-box present deliver? *Artificial Intelligence in Medicine*. 2022, vol. 124.
RASO, F. A., HILLIGOS, Hanna et al. *Artificial Intelligence & Human Rights: Opportunities & Risks*. Berkman Klein Center for Internet & Society at Harvard University, Research Publication no. 2018–6, 2018.
RIZZO, A. et al. Autonomous Virtual Human Agents for Healthcare Information Support and Clinical Interviewing. In: LUXTON, D. (ed.). *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier, 2016, pp. 53–79.
SOMASHEKHAR, S. P. et al. Early experience with IBM Watson for Oncology (WFO) cognitive computing system for lung and colorectal cancer treatment. *Journal of Clinical Oncology*. 2017, vol. 35, pp. 8527–8527.
SPARROW, R., SPARROW, L. In the hands of machines? The future of aged care. *Minds and Machines*. 2006, vol. 16, no. 2, pp. 141–161.
SUTTON, R. T., PINCOCK, D., BAUMGART, D. C. et al. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digital Medicine*. 2020, vol. 3, art. no. 17.
VALLOR, S. Carebots and Caregivers: Sustaining the Ethical Ideal of Care in the Twenty-First Century. *Philosophy & Technology.* 2011, vol. 24, art. no. 251.
WACHTER, S., MITTELSTADT, B., RUSSELL, Ch. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology*. 2018, vol. 31, no. 2. pp. 841–887.
WYNSBERGHE, A. *Healthcare Robots. Ethics, Design and Implementation*. Routledge: London, 2015, 166 p.
WYNSBERGHE, A. When Should We Use Care Robots? The Nature-of-Activities Approach. *Science and Engineering Ethics*. 2016, vol. 22, no. 6, pp. 1745–1760.

European Court of Human Rights decisions

Klass and Others v. Germany, European Court of Human Rights, 1978.
Roman Zakharov v. Russia, European Court of Human Rights, 2015.
Cevat Özel v. Turkey, European Court of Human Rights, 2016.

Other sources

Artificial Intelligence Failure at IBM ‚Watson for Oncology' [online]. Available at: https://icmrindia.org/casestudies/catalogue/IT%20and%20Systems/ITSY126.htm
Council of Europe treaty, 1981.
DEVLIN, H. *The Robots Helping NHS Surgeons Perform Better, Faster and For Longer* [online]. Available at: http://www.theguardian.com/society/2018/jul/04/robots-nhs-surgeons-keyhole-surgery-versius

Effectively Utilizing Publicly Accessible Social Media Data While Staying GDPR Compliant [online]. Available at: https://www.finextra.com/blogposting/16206/effectively-utilizing-publicly-accessible-social-media-data-while-staying-gdpr-compliant

European Convention on Human Rights.

MONTE, R. *The Beginner's Guide to an Electronic Data Capture (EDC) System* [online]. Available at: https://www.advarra.com/blog/beginners-guide-electronic-data-capture-edc-system/

MUKHERJEE, S. *A.I. Versus M.D. What happens when diagnosis is automated?* [online]. Available at: https://www.newyorker.com/magazine/2017/04/03/ai-versus-md

O'NEILL, M. *Explainer: what is artificial intelligence?* [online]. Available at: http://www.abc.net.au/news/2017-08-07/explainer-what-is-artificial-intelligence/8771632

Proposal for a regulation of the European Parliament and of the council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts (COM/2021/206).

Regulation (EU) 2016/679 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), as amended.

Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017, on medical devices, as amended.

Resolution 2011–2013, Commission Nationale de l'Informatique et des Libertés [online]. Available at: https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000024583206/