DANIELE CHIAPPINI

LEGAL AND ETHICAL STATE OF THE ART OF ARTIFICIAL INTELLIGENCE IN MEDICINE

ABSTRACT: This essay will give an overview of the state of the art on what Artificial Intelligence (AI) is, how it is already applied in the healthcare sector and the ethical and legal issues related to AI in the above-mentioned sector. Starting from the explanation of what AI is and the machine learning use in medicine, we give an overview of the AI application in medicine with a specific focus on the use of Artificial Intelligence in the COVID-19 fight. The second part of the paper examines some of the primary ethical challenges, such as human rights and access to AI, confidentiality, informed consent, safety, transparency, unfairness, biases, privacy, and data governance. Lastly, we pose an analysis of the main legal challenges and in particular civil liability, implementation, privacy, and data protection.

1. Introduction

During the last years, Artificial Intelligence applications have been used across many sectors, and the healthcare industry is one of these. The capability of Artificial Intelligence (AI) to improve outcomes in such an important subject is very intriguing and this technology has the potential to transform not only several aspects of patient care, but also the role of the physicians themselves, and can revolutionize the practice of medicine[1].

Artificial Intelligence in healthcare uses algorithms, software and neural networks to analyse data and facilitate clinical decisions.

From a historical point of view, Mycin[2], developed in the early 1970s, is one of the first examples of artificial intelligence applied in medicine with the aim to identify bacteria causing severe infections and recommend the related antibiotic therapy.

In the 1980s and 1990s, thanks to the intensified use of computers, there was an increasing development of AI systems in healthcare involving fuzzy set theory[3], Bayesian networks[45], and artificial neural networks[6].

The use of AI applications furtherly increased during the last years and today, with an estimated value of about 150 billion dollars[7], AI can be applied to robot-assisted surgery, to virtual nursing

---

1 E. Loh, Medicine and the rise of the robots a qualitative review of recent advances of artificial intelligence, Health BMJ Leader, 2018, vol. 2, p. 59-63.

2 W.J. Clancey, E.H. Shortliffe, Readings in medical artificial intelligence: the first decade, Addison-Wesley Longman Publishing Co., Inc, 1984.

3 K.P. Adlassnig, A fuzzy logical model of computer-assisted medical diagnosis, Methods of Information in Medicine, 1980, vol. 19, n. 3, p. 141.

4 A Bayesian network is a graphical model of variables and their dependencies on one another. This is used is in programs designed to compute the probability of given diseases. Symptoms can be taken as input and the probability of diseases output.

5 J.A. Reggia, Y. Peng, Modeling diagnostic reasoning: a summary of parsimonious covering theory, Computer Methods and Programs in Biomedicine, 25 (2), 1987, p. 125-134.

6 W.G. Baxt, Use of an artificial neural network for the diagnosis of myocardial infarction, Annals of Internal Medicine. 1991, 115 (11), p. 843-848.

assistants, administrative workflow, fraud detection and to reduce drugs dosage errors, helping, moreover, in both preliminary diagnoses and image diagnoses.

Such versatility is linked to a full spectrum of legal issues, from philosophy of law to human rights, privacy, contract law, labour law etc[8]. In this paper, we will give an overview of AI use in medicine and some of the related legal and ethical problems connected to AI use in healthcare.

2. Machine learning in medicine

The advantages of AI in healthcare have been extensively discussed in literature[9,10] and today copious AI-based technologies are implemented into clinical practice.

Machine learning provides systems with the ability to automatically learn and improve from experience without being explicitly programmed[11]. The process of learning begins with observations of data, direct experience, or instructions.

A machine learning-based AI uses algorithms to extract features from data, such as data derived from diagnostic imaging, genetic data and similar ones. The inputs can include patients' data (including personal data[12] and disease-specific data) and/or patient's medical outcomes, but also experiences or observations made directly by the algorithms.

Basing on the method used to process data, machine learning algorithms can be divided into different categories. The three most used are supervised learning, unsupervised learning and reinforced learning.

Supervised learning allows creating a mapping between a set of input variables and an output variable and applying this mapping to predict the outputs for unseen data[13]. Thanks to this, the algorithm is able to apply what previously learned to new data using labelled examples. Starting from the analysis of a training dataset, such as a medical data bank, the algorithm produces an inferred function to make predictions about the output. The system is able to provide output for any new input after sufficient training. The learning algorithm can also compare its output with the correct, intended output and find errors in order to modify the model accordingly[14].

On the contrary, unsupervised learning is a machine learning technique in which the machine receives inputs but obtain neither supervised target outputs nor rewards from its environment[15], using

---

7 S. Chishti, I. Bartoletti, A. Leslie, S.M. Millie, The AI Book: The artificial intelligence handbook for investors, entrepreneurs and FinTech visionaries, Wiley & sons inc., 2020, p. 79.

8 M. Parc, M. Ozer, J. Hojnik, Social and juristic challenges of artificial intelligence, Palgrave communications, 5, 2019.

9 V. Patel, E. Shortliffe, M. Stefanelli et al., The coming of age of artificial intelligence in medicine, Artificial Intelligence in Medicine 2009, vol. 46, p. 5-17.

10 Eugenio Santoro, L'intelligenza artificiale in medicina: quali limiti, quali ostacoli, quali domande, Il pensiero scientifico editore, 2017.

11 https://deepai.org/machine-learning-glossary-and-terms/machine-learning, last visit January 2021.

12 In the General Data Protection Regulation (REGULATION EU 679/2016 - GDPR) personal data means any information relating to an identified or identifiable natural person. An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

13 P. Cunningham, M. Cord, S.J. Delany, Supervised Learning, Machine Learning Techniques for Multimedia, Cognitive Technologies, Springer, 2008, p. 21.

14 P. Ganapathi, Shanmugapriya D., Handbook of Research on Machine and Deep Learning Applications for Cyber Security, Advances in Information Security, Privacy, and Ethics, IGI Global, 2019, p. 383.

15 Z. Ghahramani, Unsupervised Learning In: Bousquet O., von Luxburg U., Rätsch G. (eds) Advanced Lectures on Machine Learning, Lecture Notes in Computer Science, vol 3176. Springer, 2004, p. 73.

information that is not classified or labelled. These algorithms study how systems can infer a function to describe a hidden structure from unlabelled data[16].

Recently has been introduced semisupervised learning as a hybrid between the two above mentioned processes.

Reinforced learning is instead a method in which the machine interacts with its environment by producing actions that affect the state of the environment which, in turn, result in the machine receiving scalar rewards or punishments[17]. This method allows machines and software agents to automatically establish the correct behaviour within a specific context in order to maximize performance[18].

Among the presented processes, supervised learning provides more clinically relevant results[19]; hence AI applications in healthcare most often use supervised learning. Relevant techniques include linear regression, logistic regression, naïve Bayes, decision tree, nearest neighbour, random forest, discriminant analysis, support vector machine (SVM) and neural network. In particular, it has been investigated that about one-third of the machine learning algorithms used in the medical literature belongs to the Support Vector Machine technique and another third belongs to Neural Network[20] technique.

Support Vector Machine are particular linear classifiers used for classifying the subjects into different groups. The SVM accomplishes the classification task by constructing, in a higher-dimensional space, the hyperplane that optimally separates the data into N (often two) categories.[21] The outcome represents whether the patient belongs to group 1 or to the group N according to a decision based on a specific trait. SVM has been extensively used in medical research[22,23].

Neural network is a computational model, available since the 1960s, inspired by biological neural networks, like the ones in the human brain. Thanks to a series of functions, it processes an input signal and translates it over several stages into the output. This technology is well established in healthcare research for several decades[24].

Connected to neural network is the concept of deep learning. In a 2016 Google Tech Talk, Jeff Dean describes deep learning algorithms as using very deep neural networks, where "deep" refers to the number of layers or iterations between input and output[25]. The structure of neural networks, organized in multiple layers, allows them to address more complex tasks. This, in several clinical cases, demonstrated superior performance as compared with classical machine learning algorithms and, in

---

[16] P. Ganapathi, Shanmugapriya D., Handbook of Research on Machine and Deep Learning Applications for Cyber Security, Advances in Information Security, Privacy, and Ethics, IGI Global, 2019, p. 383.

[17] Z. Ghahramani, Unsupervised Learning In: Bousquet O., von Luxburg U., Rätsch G. (eds) Advanced Lectures on Machine Learning, Lecture Notes in Computer Science, vol 3176. Springer, 2004, p. 73.

[18] J. Lentin, ROS robotic project, Packt publishing, 2017, p. 203.

[19] F. Jiang, Y. Jiang, H. Zhi et al., Artificial intelligence in healthcare: past, present and future, Stroke and Vascular Neurology, 2017, vol 2, p.234.

[20] F. Jiang, Y. Jiang, H. Zhi et al., Artificial intelligence in healthcare: past, present and future, Stroke and Vascular Neurology, 2017, vol 2, figure 5, p. 234. The data used in the article have been generated through searching the machine learning algorithms within healthcare on PubMed.

[21] M. Adankon, M. Cheriet, Support Vector Machine, In: Li S.Z., Jain A. (eds) Encyclopedia of Biometrics, Springer, 2009.

[22] G. Orrù, W. Pettersson-Yeo, A.F. Marquand et al., Using support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review, Neurosci Biobehav Rev 2012, 36, p. 1140-1152.

[23] N.H. Sweilam, A.A. Tharwat, N.K.A. Moniem, Support vector machine for diagnosis Cancer disease: a comparative study, Egyptian Informatics Journal 2010, 11, p. 81–92.

[24] T. Davenport, R. Kalakota, The potential for artificial intelligence in healthcare, Future Healthcare Journal, 6(2), 2019, p. 94.

[25] https://deepai.org/machine-learning-glossary-and-terms/machine-learning, last visit January 2021.

some cases, achieved comparable or better performance than clinical experts[26]. The current trend for deep learning in medicine described by a recent article[27] shows that from 2013 to 2016 its application in the medical research field increased ten times.

3. AI application in healthcare

The first applications of AI in medicine have been developed in the 1960s and 1970s. The first problem-solving program, or expert system, known as Dendral[28], actually designed for applications in organic chemistry, provided the basis for the Mycin system, considered one of the most significant early uses of artificial intelligence in medicine. Since that moment, the use of Artificial Intelligence in healthcare has grown exponentially.

Mapping exhaustively the current Artificial Intelligence applications in healthcare is quite difficult, but a quick overview of its main uses and developments will help to give an idea about how AI is currently utilized in the different medical fields.

3.1. Disease Diagnosis

Many Artificial Intelligence algorithms have been used to efficiently and accurately diagnose diseases, even if the research mainly concentrates only on few disease types: cancer, nervous system disease and cardiovascular disease[29]. According to a recent study[30], Artificial Intelligence may be able to diagnose diseases as successfully as human healthcare physicians. The study compared the performance of deep learning AI and healthcare professionals in detecting diseases from medical imaging, using studies carried out between 2012 and 2019. As a result, the researchers found that, in the past few years, AI has become more accurate and more viable source of diagnostic information. According to the study, the diagnostic performances were found to be equivalent, and deep learning may become even more efficient in identifying diagnosis in the next years.

Within the different types of AI, a comprehensive study[31] reviewed papers wrote within 2008-2017, concluding that the Artificial Neural Network is able to classify diabetes and cardiovascular diseases more accurately respect the Bayesian Network.

Moreover, pharmaceutical companies recently started working with tech companies to create software to help diagnose complex and rare conditions and consequently help develop new drugs to treat these illnesses[32].

---

26 G. Chassagnon, M. Vakalopolou, N. Paragios et al., Deep learning: definition and perspectives for thoracic imaging, European Radiology 30, 2020, p. 2021-2030.

27 F. Jiang, Y. Jiang, H. Zhi et al., Artificial intelligence in healthcare: past, present and future, Stroke and Vascular Neurology, 2017, vol 2, figure 10, p. 238. The data used in the article have been generated through searching the deep learning in healthcare and disease category on PubMed.

28 R.K. Lindsay, B.G. Buchanan, E.A. Feigenbaum, J. Lederberg, Dendral: a case study of the first expert system for scientific hypothesis formation, Artificial Intelligence, 61 (2), 1993, p. 209.

29 F. Jiang, Y. Jiang, H. Zhi et al., Artificial intelligence in healthcare: past, present and future, Stroke and Vascular Neurology, 2017, vol 2, p. 231.

30 X. Liu*, L. Faes*, A.U. Kale, S.K. Wagner, D.J. Fu, A. Bruynseels, T. Mahendiran, G. Moraes, M. Shamdas, C. Kern, J.R. Ledsam, M.K. Schmid, K. Balaskas, E.J. Topol, L.M. Bachmann, P.A. Keane, A.K. Denniston, A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis, Lancet Digital Health 2019, 1, p. 271.

31 B. Alić, L. Gurbeta and A. Badnjević, Machine learning techniques for classification of diabetes and cardiovascular diseases, 6th Mediterranean Conference on Embedded Computing (MECO), 2017, pp. 1-4

3.2. Drugs creation and interaction

Artificial Intelligence already created a drug used to treat OCD (obsessive-compulsive disorder). The molecule DSP-1181 was invented by an artificial intelligence through joint efforts of Exscientia and Sumitomo Dainippon Pharma and it entered in Phase I clinical trial in Japan[33]. The drug development took a single year, while pharmaceutical companies usually spend about five years on similar projects. Moreover, the development of a new drug is a very expensive process and it seems that the use of AI could decrease the research costs.

This is the first clinical trial based on an experimental molecule created by AI, but it is only the first step, as AI is currently applied in other experimentations. Pfizer, for example, is using IBM Watson to power its search for immuno-oncology drugs. Sanofi has signed a deal to use Exscientia's artificial-intelligence to research metabolic-disease therapies, Roche subsidiary Genentech is using an Artificial Intelligence system from GNS Healthcare to help the multinational company's search for cancer treatments[34] and other biopharma industries have similar collaborations or internal programmes involving AI.

Artificial intelligence can also be used to predict the side effects of a drug, as shown by a Stanford University study[35]. According to the study, the algorithm, starting from the chemical structure of the potential drug, is able to make predictions both on the possible toxicity and on the instability of the molecule, significantly speeding up the development process.

Another aspect is related to the development of algorithms to identify drug-drug interactions to discover the existence of threats for those taking multiple medications simultaneously. Machine learning algorithms using natural language processing can extract information on interacting drugs and their possible effects from medical literature as shown by the DDI corpus[36].

Other algorithms are able to identify drug-drug interactions starting from user-generated content, especially from electronic health records and/or adverse event reports[37]. Other sources of information are represented by databases such as the FDA Adverse Event Reporting System (FAERS) and the World Health Organization's VigiBase that allow doctors to submit reports of possible negative reactions to medications. Starting from these data, deep learning algorithms have been trained to parse these reports and detect patterns that imply drug-drug interactions[38].

3.3. Medical Imaging

---

32 Bayer and Merck have recently received FDA breakthrough designation to artificial intelligence software for CTEPH pattern recognition. This new software will be deployed through Bayer's Radimetrics platform detecting a rare but severe disease and help save and improve the lives of patients. https://pharma.bayer.com/artificial-intelligence-when-we-suddenly-know-what-we-dont-know, last visit July 2020.

33 T. Burki, A new paradigm for drug development, Lancet Digital Health, 2020.

34 https://www.nature.com/articles/d41586-018-05267-x, last visit January 2021.

35 H. Altae-Tran et al., Low Data Drug Discovery with One-Shot Learning, ACS Central Science, http://pubs.acs.org/doi/abs/10.1021/acscentsci.6b00367, last visit January 2021.

36 M. Herrero-Zazo, I Segura-Bedmar, P. Martínez, T. Declerck, The DDI corpus: An annotated corpus with pharmacological substances and drug-drug interactions, Journal of Biomedical Informatics, Volume 46, Issue 5, 2013, p. 914-920.

37 F. Christopoulou, T.T. Tran, S.K. Sahu, M. Miwa, S. Ananiadou, Adverse drug events and medication relation extraction in electronic health records with ensemble deep learning methods. Journal of the American Medical Informatics Association, 2020, 27(1), p. 39-46.

38 B. Xu et al., Incorporating User Generated Content for Drug Drug Interaction Extraction Based on Full Attention Mechanism, in IEEE Transactions on NanoBioscience, 2019, vol. 18, no. 3, pp. 360-367.

Artificial Intelligence (AI) is applied to medical imaging to help patient diagnosis. Medical imaging uses different processes and imaging methods to represent an internal image of the human body for diagnostic and treatment purposes. It includes various radiological imaging techniques such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI), medical ultrasonography or ultrasound, Positron Emission Tomography (PET) and X-ray radiography.

The increasing amount of data can influence how radiologists interpret images since when too much time is taken for image analysis, the time for evaluating clinical and laboratory contexts is squeezed[39]. So, AI can improve traditional medical imaging by offering computational capabilities that process images with greater speed and accuracy.

Current AI applications for medical imaging include the diagnose of neurological conditions like Amyotrophic Lateral Sclerosis (ALS) or to predict Alzheimer's disease onset years before it manifests[40]. Some applications are able to reveal cardiovascular abnormalities and indicate the risk of correlated cardiovascular disease[41] while others can make an early cancer diagnosis using an AI based on Convolutional Neural Networks to identify various types of cancer[42].

About the success rate, in 2018, a paper[43] determined that skin cancer can be detected more accurately by artificial intelligence, based on deep learning convolutional neural network, than by dermatologists, causing concern among the specialists[44].

3.4. Mental healthcare

In the mental health field, AI applications are being adopted more slowly than in other healthcare sectors[45]. Despite this, it is sustained that Artificial Intelligence has great potential to help in the diagnosis and understanding of mental illnesses[46] and it is already adopted in studies for anxiety and depression[47] and in medical chatbots development[48]. Accordingly, the use of "AI" and "Mental health" terms in publications increased in the last five years and its use is changing psychiatry[49].

---

[39] F. Pesapane, M. Codari, F. Sardanelli, Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine, European Radiology Experimental, 2018, 2, p. 4.

[40] Y. Ding, H. Ho Sohn, M.G. Kawczynski et al., A Deep Learning Model to Predict a Diagnosis of Alzheimer Disease by Using 18F-FDG PET of the Brain, Radiology 2019, 290:2, p. 456-464.

[41] K.R. Siegersma, T. Leiner, D.P. Chew et al., Artificial intelligence in cardiovascular imaging: state of the art and implications for the imaging cardiologist, Netherlands Heart Journal, 2019, 27, p. 403-413.

[42] S. Dabeer, M.M. Khan, S. Islam, Cancer diagnosis in histopathological image: CNN based approach, Informatics in Medicine Unlocked, Volume 16, 2019, https://www.sciencedirect.com/science/article/pii/S2352914819301133, last visit January 2021.

[43] H.A. Haenssle, C. Fink, F. Toberer et al., Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions, Annals of Oncology, 2020, https://www.annalsofoncology.org/action/showPdf?pii=S0923-7534%2819%2935468-7, last visit January 2021.

[44] K. Chockley, E. Emanuel, The End of Radiology? Three Threats to the Future Practice of Radiology, Journal of the American College of Radiology, 2016, https://www.jacr.org/action/showPdf?pii=S1546-1440%2816%2930590-7, last visit January 2021.

[45] S. Graham, C. Depp, E. Lee, C. Nebeker, X. Tu, H. Kim, D. Jeste, Artificial Intelligence for Mental Health and Mental Illnesses: An Overview. Current Psychiatry Reports. Springer, 2019.

[46] D. Bzdok, A. Meyer-Lindenberg, Machine Learning for Precision Psychiatry: Opportunities and Challenges, Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, Volume 3, Issue 3, 2018, p. 223-230.

[47] R. Fulmer, A. Joerin, B. Gentile, L. Lakerink, M. Rauws, Using Psychological Artificial Intelligence (Tess) to Relieve Symptoms of Depression and Anxiety: Randomized Controlled Trial, JMIR Mental Health, 2018, 5(4).

[48] Artificial conversational agents able to imitate human behaviour.

[49] M. Brunn, A. Diefenbacher P. Courtet et al. The Future is Knocking: How Artificial Intelligence Will Fundamentally Change Psychiatry, Springer, 2020, Academic Psychiatry 44, p. 461-466.

According to the literature[50], leverage AI techniques offers the ability to develop better prediagnosis screening tools and formulate risk models to determine individual predisposition or risk of developing mental illness. The importance of these applications is also proved by the fact that also private corporations are interested in, such as the screening for suicidal ideation implemented by Facebook in 2017 demonstrate [51].

3.5. Industries and AI in healthcare

During the last years, a high number of ICT companies studied AI implementation in several everyday life aspects, included the healthcare sector. As the number of data increases, AI decision support systems become more efficient and, consequently, numerous companies are exploring the possibilities of its incorporation in the healthcare industry[52]. This, according to the yearly reports "Artificial Intelligence in Medicine Market - Growth, Trends, and Forecast", will increase, only in the medical sector, the artificial intelligence market value from 2,24 billion dollars in 2018 to 23,85 billion dollars by 2025[53]. The following examples are related to the major corporations' activities in AI and the healthcare field.

3.5.1. Google

Google created Google Health in 2006, with the aim to store health records and data in order to directly connect doctors, hospitals and pharmacies. The programme was closed in 2012, but in 2014 Google acquired DeepMind and in 2016 announced a collaboration to develop Artificial Intelligence applications for healthcare, Google DeepMind Health. In 2018, Google combined the Google Research health team, DeepMind Health and a team from Google's Hardware division focused on health-related applications, to form Google Health, publishing research papers and building tools in collaboration with a variety of healthcare partners. According to the Google AI blog[54], the following are some highlights from 2019:

Firstly, Google researchers identified a deep learning model for mammography that can assist physicians in spotting breast cancer[55] with greater accuracy, reducing both false positives and false negatives. The model trained on de-identified data from a UK hospital showed similar results in accuracy when used to evaluate patients in a completely different healthcare system in the U.S.

Another result concerns a deep learning model for differential diagnoses of skin diseases[56] that can give results significantly more accurate than the primary care physicians ones and on par with or perhaps slightly better than dermatologists.

Moreover, working alongside experts from the US Department of Veterans Affairs (VA), Google Health researchers showed that a machine learning model can predict the onset of an acute kidney

---

[50] S. Graham, C. Depp, E. Lee, C. Nebeker, X. Tu, H. Kim, D. Jeste, Artificial Intelligence for Mental Health and Mental Illnesses: An Overview. Current Psychiatry Reports. Springer, 2019.

[51] G. Coppersmith, R. Leary, P. Crutchley, A. Fine, Natural Language Processing of Social Media as Screening for Suicide Risk, Biomedical Informatics Insights, 2018, https://journals.sagepub.com/doi/10.1177/1178222618792860#articleCitationDownloadContainer, last visit January 2021.

[52] D. Arnold, T. Wilson, What Doctor? Why AI and robotics will define New Health, PwC, 2017.

[53] Research and markets, Artificial Intelligence in Medicine Market - Growth, Trends, and Forecast, reports 2019-2024 and 2020-2025.

[54] Google AI blog, the latest news from Google AI, https://ai.googleblog.com/2020/01/google-research-looking-back-at-2019.html, last visit January 2021.

[55] https://blog.google/technology/health/improving-breast-cancer-screening, last visit January 2021.

[56] https://ai.googleblog.com/2019/09/using-deep-learning-to-inform.html, last visit January 2021.

injury[57] (AKI), one of the leading causes of avoidable patient harm, up to two days before it happens. This could give doctors a 48-hour head start in treating this serious condition.

About other subject, it has been developed a deep learning model to predict lung cancer[58] by examining the results of a single CT scan. The study showed that the model outputs are par or better than trained radiologists ones at early detection of lung cancer.

In addition, Google continued to expand and evaluate machine learning tools for detection and prevention of eye disease[59], in collaboration with Verily and other partners. It has been published a research paper on an augmented reality microscope for cancer diagnosis, whereby pathologists can get real-time feedback about what parts of a slide are most interesting while examining tissue through a microscope.

3.5.2. IBM

IBM, in 2011, created Watson, a question-answering computer system capable of answering questions posed in natural language. In 2013 Watson software has been adopted to provide utilization management decisions in lung cancer treatment at Memorial Sloan Kettering Cancer Centre.

Consequently, IBM funded the Watson Health division in 2015, and in 2016 announced, in collaboration with the Manipal Hospital, the launch of IMB Watson for Oncology[60], currently in development at Memorial Sloan Kettering Cancer Centre and Cleveland Clinic. Moreover, in May 2017, IBM and Rensselaer Polytechnic Institute began a joint project entitled Health Empowerment by Analytics, Learning and Semantics (HEALS), to explore using AI technology to enhance healthcare[61].

With IBM Watson Health the company intends to create software able to help health professionals and researchers translate data and knowledge into insights to make more informed decisions about care[62].

Since the decision to bring Watson into the health industry, the company has made nearly 50 announcements about partnerships intended to develop new AI-enabled tools for medicine. Unfortunately, many of these alliances have not yet led to commercial products, however, IBM sustains that research efforts have been valuable and that many cooperations are currently ongoing. Actually, according to a recent study[63], in 24 launched projects, only five developed an output tool.

3.5.3. Microsoft

Microsoft affirms that is currently "using AI to unearth new insights, advance discoveries, and explore solutions with partners to improve health outcomes globally"[64].

The company uses AI applications in different fields, like precision medicine, in which Artificial Intelligence enables new personalized treatments by taking into consideration aspects like patients' genetics and lifestyle to radically improve health and longevity[65].

---

57 https://www.nature.com/articles/s41586-019-1390-1, last visit January 2021.
58 https://www.blog.google/technology/health/lung-cancer-prediction/, last visit January 2021.
59 https://www.blog.google/technology/health/new-milestones-helping-prevent-eye-disease-verily/, last visit January 2021.
60 https://www.ibm.com/watson-health/oncology-and-genomics, last visit January 2021.
61 https://news.rpi.edu/content/2017/05/17/ibm-and-rensselaer-team-research-chronic-diseases-cognitive-computing, last visit January 2021.
62 https://www.ibm.com/watson-health/about/get-the-facts, last visit January 2021.
63 E. Strickland, How IBM Watson Overpromised and Underdelivered on AI Health Care, IEEE spectrum, 2019.
64 https://www.microsoft.com/en-us/ai/innovation-in-health, last visit January 2021.
65 https://www.microsoft.com/en-us/research/project/immunomics/, last visit January 2021.

Another aspect in which Microsoft is using AI is the population health analytics. Using Artificial Intelligence, Microsoft intends to analyse data from social determinants of health to environmental issues and disease vectors to better predict problems before they occur or spread[66].

Moreover, with the Hannover project[67], Microsoft intends to use AI to allow researchers and medical professionals to quickly filter information from all types of sources, uncover new insights and provide treatments more quickly to those in need.

The importance of AI in healthcare for Microsoft is also proved by the white paper released about the use of AI and Imaging Data in Healthcare[68].

3.5.4. Apple

Apple has a completely different approach to AI in healthcare. Instead of creating specific Artificial Intelligence programmes or applications, the company is focussed on using its devices and AI to simplify or track the health status of the customers/patients. According to the Apple website[69], apps on iPhone and Apple Watch allow clinicians to access health records and data, nurses to ensure better patient safety while administering medication and patients to stay informed and engaged in their own care by communicating with their medical teams during hospital stays. Moreover, Apple gives medical researchers access to ResearchKit, a way to build their own app. In January 2020 Apple has 1.5 billion active devices worldwide[70] most of them capable to acquire and collect health data.

3.4.6. Other companies

The previous examples are not exhaustive, since a lot of other companies are currently working on AI and health. A high number of applications use AI algorithms to give medical advice, like Babylon Health's[71], Ada Health[72], AliHealth Doctor You[73], KareXpert[74] and Healthily[75]. With these apps, users are able to report their symptoms and the app, using speech recognition, compares it with a database of illnesses.

Some companies developed robot services integrated with Artificial Intelligence technology to identify a registered customer and provide personalized recommendations in different medical areas or in the field of medical imaging. These kinds of robots have been developed by IFlytek with the robot

---

[66] https://www.microsoft.com/en-us/research/project/project-premonition/, last visit January 2021.

[67] https://www.microsoft.com/en-us/research/project/project-hanover/, last visit January 2021.

[68] A. Kowala, L. Matragrano, P. Shah, Use AI and Imaging Data to Unlock Insights and Improve Healthcare, https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/use-ai-and-imaging-data-to-unlock-insights-and-improve-healthcare-paper.pdf, last visit January 2021.

[69] https://www.apple.com/healthcare/, last visit January 2021.

[70] https://www.macrumors.com/2020/01/28/apple-1-5-billion-active-devices-worldwide/, last visit January 2021.

[71] Babylon Health through deep learning Babylon's A.I. system can provide users with personalized insights to stay healthy and better understand their health. In 2019 people used its app for a total of 2,2 million A.I. consultations. Source: https://www.babylonhealth.com/, last visit January 2021.

[72] Ada's core system connects medical knowledge with intelligent technology to help all people actively manage their health and medical professionals to deliver effective care. Ada has 10 million users and 20 million health assessments completed. Source: https:/ada.com/, last visit January 2021.

[73] Doctor You uses AI to diagnose medical images of CT scans and identify inflammatory cells in human organs as an early indicator of cancer.

[74] KareXpert is a technology Company to design patented "Advanced Health Cloud Technology Stack" which connects Patients with all the care Providers with the purpose of reinvent and redefine the Indian healthcare system. Source: https://www.karexpert.com/dhs/about-us/, last visit January 2021.

[75] Your.MD developed Healthily, a free app where anyone, anywhere can find safe and personalized information, guidance and support for their health. Source: https://www.your.md/app, last visit January 2021.

"Xiao Miao"[76] and Softbank Robotics with "Pepper" and "Nao"[77]. The value of these robots is increasing and a lot of companies are currently studying the possibilities that robotics and AI can offer. For example, a Chinese artificial intelligence-powered robot named Xiaoyi (meaning Little Doctor) is the first robot who passed the written test of China's national medical licensing examination[78].

Moreover, a high number of companies are investing in Artificial intelligence models able to make early detection of cancer, like Huiying's Breast Cancer Detection Solution[79] built using Intel's hardware or Kheiron Medical developed deep learning software to detect breast cancers in mammograms[80], while Tencent is working on several medical systems and services, like Miying[81], a system able to help gynaecologists efficiently identify cervical transformation zone, cervical intraepithelial neoplasia (CIN) and location of lesions in the cervical cancer screening process.

4. AI and COVID-19

On January 9, 2020, the World Health Organization (WHO) communicated a flu-like outbreak in China[82] following a cluster of pneumonia cases reported in Wuhan. On January 30, the WHO declared the coronavirus outbreak a Public Health Emergency of International Concern[83]. On March 11 COVID-19 has been proclaimed by the WHO Director-General as a pandemic[84].

In this global outbreak, for the first time, Artificial intelligence was being a useful tool to monitor and respond to the crisis. In the following pages, we'll give a short, not-exhaustive, overview of how Artificial intelligence and data science technologies are currently being used to fight COVID-19.

4.1. Identification, prediction and tracking outbreaks

AI is also applied in the identification, prediction and tracking of disease outbreaks. Currently, there are a lot of studies and start-ups working in this field. In particular, BlueDot[85] is a start-up that built a sophisticated AI platform to identify disease outbreaks and quantifying the risk of exposure to infectious diseases. The AI processes data obtained by different sources, such as the world's air travel network, health reports, and many others on over 150 diseases and syndromes around the world. These data include official data from national and supranational organizations and less structured information,

76 X. Kong, B. Ai,Y. Kong, L. Su et al., Artificial intelligence: a key to relieve China's insufficient and unequally-distributed medical resources, American journal of translational research, 2019, 11(5), p. 2632-2640.
77 Pepper and NAO can assist patients in self-diagnosis, support staff in health trending & monitoring. They are also the platform for telemedicine and hub of information distribution like alert, notifications, fall & sound detection. Source: https://www.softbankrobotics.com/emea/en/industries/healthcare, last visit January 2021.
78 https://cmte.ieee.org/futuredirections/2017/12/02/congrats-xiaoyi-you-are-now-a-medical-doctor/, last visit January 2021.
79 Huiying's Breast Cancer Detection Solution, https://builders.intel.com/docs/aibuilders/huiying-medical-technology-optimizes-breast-cancer-early-screening-and-diagnosis-with-intel-ai-technologies.pdf, last visit January 2021.
80 https://www.kheironmed.com/meet-mia, last visit January 2021.
81 https://www.tencent.com/en-us/articles/2200912.html, last visit January 2021.
82 https://www.who.int/china/news/detail/09-01-2020-who-statement-regarding-cluster-of-pneumonia-cases-in-wuhan-china, last visit January 2021.
83 http://www.euro.who.int/en/health-topics/health-emergencies/international-health-regulations/news/news/2020/2/2019-ncov-outbreak-is-an-emergency-of-international-concern, last visit January 2021.
84 https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020, last visit January 2021.
85 https://bluedot.global, last visit January 2021.

like media sources, in 65 languages per day. Epidemic detection examines over 150 different pathogens, toxins and syndromes in near real-time[86].

About the coronavirus outbreak, BlueDot launched its first alarm on December 30th, 2019[87], publishing a paper related to the potential spread of the virus immediately thereafter[88]. To compare, the World Health Organization, on January the 5th, published a news[89] in which it didn't recommend "any specific measures for travellers. In case of symptoms suggestive of respiratory illness either during or after travel, travellers are encouraged to seek medical attention and share travel history with their healthcare provider". Moreover, WHO advised that "against the application of any travel or trade restrictions on China based on the current information available on this event."

A similar project has been developed by the Massachusetts Institute of Technology, where a team trained a machine learning model to allow it to accurately predicts the spread of COVID-19. Using mixed first-principles epidemiological equations and data-driven neural network, the model helps to interpret and extrapolate publicly-available data for insights into the disease's spread, taking into account different factors such as social distancing, quarantine and standard epidemiology parameters handled by different governments[90].

A Harvard and MIT mixed team is using machine learning to diagnose COVID-19 by voice since the presence of COVID-19 could be detected using speech signals[91]. The AI developed by the scientists uses voice recordings from COVID-19 patients and healthy people as training data to identify specific vocal signatures that could indicate the presence of an infection. Moreover, at the Swiss university École Polytechnique Fédérale de Lausanne, a team is building an app called "Coughvid"[92] able to listen to people cough and use AI to determine whether they sound like a coronavirus patient, using as data audio samples of people coughing.

4.2. Diagnosis, precaution and protection

Artificial Intelligence is also used in coronavirus diagnosis. Chinese Alibaba DAMO academy, the research branch of the Alibaba company, sustains that an AI has been trained to detect Covid-19 using as a dataset more than 5,000 confirmed cases. According to the company, the system is able to process the 300 to 400 scans needed to diagnose a coronavirus in 20 to 30 seconds with an accuracy of 96%[93]. The system seems to have helped at least 26 Chinese hospitals to review more than 30,000 cases[94].

Another example is provided by a South Korean company that started to develop a testing kit as soon as the virus has been identified. They used artificial intelligence to reduce the time needed to

---

[86] https://bluedot.global, how it works. Last visit January 2021.
[87] Z. Allam, G. Dey, D.S. Jones, Artificial Intelligence (AI) Provided Early Detection of the Coronavirus (COVID-19) in China and Will Influence Future Urban Health Policy Internationally, AI 2020, 1, p. 156-165.
[88] I.I. Bogoch, A. Watts, A. Thomas-Bachli, C. Huber, M.U.G. Kraemer, K. Khan, Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel, Journal of Travel Medicine, Volume 27, Issue 2, March 2020.
[89] https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/, last visit January 2021.
[90] R. Dandekar, G. Barbastathis, Quantifying the effect of quarantine control in Covid-19 infectious spread using machine learning, MedRxiv 2020. Note that this is a preprint.
[91] https://www.ll.mit.edu/news/signs-covid-19-may-be-hidden-speech-signals, last visit January 2021.
[92] The coughvid developer "propose to leverage signal processing, pervasive computing, and machine learning to develop an Android application and website to automatically screen COVID-19 from the comfort of people's homes. Test subjects will be able to simply download a mobile application, enter their symptoms, record an audio clip of their cough, and upload the data anonymously to our servers. We will then use audio signal processing and machine learning techniques to evaluate if there is some room for automatic or assisted COVID-19 screening". https://coughvid.epfl.ch/about/, last visit January 2021.
[93] https://www.coe.int/en/web/artificial-intelligence/ai-and-control-of-covid-19-coronavirus. Last visit January 2021.
[94] https://www.alizila.com/how-damo-academys-ai-system-detects-coronavirus-cases/. Last visit January 2021.

design testing kits based on the genetic make-up of the virus. Consequently, 118 medical facilities were able to use the device and more than 230,000 people have been tested[95] during the first wave of the outbreak.

4.3. Cure and drugs development

This is a particular case of the wider AI use for drugs development (see chapter 3.2). In this case, the AI can assist researchers to find a vaccine and contain the pandemic.

In particular, AI provided significant support in predictions of the virus structure. The American start-up Moderna has distinguished itself by its mastery of a biotechnology-based on messenger ribonucleic acid (mRNA) for which the study of protein folding is essential. It has managed to significantly reduce the time required to develop a prototype vaccine testable on humans thanks to the support of bioinformatics, of which AI is an integral part[96].

Similarly, the Chinese company Baidu, in partnership with Oregon State University and the University of Rochester, published its Linearfold prediction algorithm[97]. This is an algorithm much faster than traditional algorithms in predicting the structure of a virus' secondary ribonucleic acid (RNA). Moreover, it provides scientists with additional information on how the virus spread. The prediction of the secondary structure of the RNA sequence of Covid-19 would thus have been calculated by Linearfold in 27 seconds instead of 55 minutes[98].

ChemRxiv uses Artificial Intelligence algorithms to identify "progeny" drugs that are similar to the "parents" already being tested against COVID-19. These algorithms assess similarity not only by the molecular make-up of the molecules but also by the "context" in which specific functional groups are arranged and/or by the three-dimensional distribution of pharmacophores[99]. Progenies are essentially either already approved drugs or are medications in advanced clinical trials. However, should the currently tested parent medicines fail, the progenies could be re-purposed against the COVID-19.

Other companies started studying the virus structure with AI. Google DeepMind Alphabet, for example, shared its predictions of coronavirus protein structures with its AlphaFold AI system[100], while IBM, Amazon, Google and Microsoft provided the computing power of their servers to the US authorities to process very large datasets in epidemiology, bioinformatics and molecular modelling[101].

4.4. Identification and risk predictions of the infected

A Zhongnan Hospital team at Wuhan (China) is using an AI software model to detect COVID-19 by visual signs[102]. The model helped the Chinese staff in screening and prioritization patients that are

---

[95] https://edition.cnn.com/2020/03/12/asia/coronavirus-south-korea-testing-intl-hnk/index.html, last visit January 2021.
[96] https://www.coe.int/en/web/artificial-intelligence/ai-and-control-of-covid-19-coronavirus, last visit January 2021.
[97] L. Huang, H. Zhang, D. Deng, K. Zhao, K. Liu, D.A. Hendrix, D.H. Mathews, LinearFold: linear-time approximate RNA folding by 5'-to-3' dynamic programming and beam search, Bioinformatics, Volume 35, Issue 14, 2019, Pages i295-i304.
[98] https://www.technologyreview.com/s/615342/how-baidu-is-bringing-ai-to-the-fight-against-coronavirus/, last visit January 2021.
[99] M. Moskal, W. Beker, R. Roszak, E.P. Gajewska, A. Wołos, K, Molga et al., Suggestions for second-pass anti-COVID-19 drugs based on the Artificial Intelligence measures of molecular similarity, shape and pharmacophore distribution, ChemRxiv, 2020. Note this is a preprint.
[100] https://deepmind.com/research/open-source/computational-predictions-of-protein-structures-associated-with-COVID-19. Last visit January 2021.
[101] https://techcrunch-com.cdn.ampproject.org/c/s/techcrunch.com/2020/03/22/ibm-amazon-google-and-microsoft-partner-with-white-house-to-provide-compute-resources-for-covid-19-research/amp/. Last visit January 2021.
[102] https://www.itnonline.com/content/infervision-launches-solution-fight-against-coronavirus, last visit January 2021.

likely to have the virus. Developed by Infervision[103] startup, the software was developed to detect lung cancer using CT, but now it has been trained, with more than 22,000 data, to detect signs of pneumonia caused by the virus. The algorithm is able to localise pneumonia despite the fact that images origin from different hospitals and were considerably varying in technique, contrast and resolution. Until the 28th of February[104], over 34 hospitals in China used this AI and successfully detect more than 32,000 COVID-19 cases. Moreover, thanks to an agreement between the European Commission and Infervision, the algorithm is in the experimentation phase at Turin Health City[105].

Computer scientists at the University of Copenhagen are currently developing an AI model capable of calculating the probability of an individual patient's need for a ventilator or intensive care[106]. The model is not aimed to treat individual patients, but its objective is to create a tool for hospitals to plan and deploy their resources in the most efficient way.

4.5. AI and population control

AI has been used in support of mass surveillance policies in China, to measure temperature and recognize individuals[107]. It is interesting to underline that facial recognition devices have experienced difficulties due to the wearing of surgical masks. Because of this, companies are attempting to circumvent this problem since many services in China rely on this technology, including state services for surveillance measures. Hanvon (Hanwang Technology Ltd) sustains to have created a technology that can successfully recognize people who wear masks with a recognition rate of about 95%[108].

Moreover, GAFAM[109] members have at their disposal information extremely valuable in case of crisis: a huge amount of different data related to the worldwide population. The U.S. government asked these companies to have access to American user's data, especially about mobile phones data, in order to fight the spread of the virus[110], however, this approach could lead to legal risk and potential image damage to the companies. For example, Google is resistant to use the location data to track and identify individuals, and the Google Health's Head of Communications and Public Affairs, Johnny Luu, said that Google was "exploring ways that aggregated anonymized location information could help in the fight against COVID-19"[111]. Anyway, it is clear that the data gathered by the ICT companies are fundamental to train Artificial Intelligence also in the response to the coronavirus outbreak, but all this amount of data can be used for shadowy aims, in particular in non-democratic countries.

4.6 Fake news spread

---

[103] https://global.infervision.com, Last visit January 2021.
[104] https://ec.europa.eu/futurium/en/european-ai-alliance/how-artificial-intelligence-aiding-fight-against-coronavirus, last visit January 2021.
[105] https://www.unito.it/comunicati_stampa/torino-arriva-lintelligenza-artificiale-combattere-il-covid-19, last visit January 2021.
[106] https://www.eurekalert.org/pub_releases/2020-04/uoc-ait040220.php, last visit January 2021.
[107] According to KC Wearable's global chief, Dr. Jie Guo, more than 1,000 "smart helmets" capable of flagging individuals with high body temperature are used by Chinese law enforcement. The company says that helmets have been sent to Italy's carabinieri military police and to the Dutch government for testing. Police are also using the devices in Dubai. The Business Insider magazine approached the Italian embassy and the Dutch government for comment but neither responded.
[108] M. Pollard, Even mask-wearers can be ID'd, China facial recognition firm says, Reuters, 9 March 2020.
[109] GAFAM is an acronym for Google, Apple, Facebook, Amazon and Microsoft companies.
[110] T. Romm, E. Dwoskin, C. Timberg, U.S. government, tech industry discussing ways to use smartphone location data to combat coronavirus, The Washington Post, March 18, 2020.
[111] https://www.darkdaily.com/tech-companies-suggest-ways-location-tracking-could-help-health-authorities-fight-the-covid-19-coronavirus-by-identifying-people-who-may-need-clinical-laboratory-testing/, last visit January 2021.

A particular aspect of AI use is related to the fake news spread during the first wave of the coronavirus pandemic. A study published in the Boston MIT magazine[112] identified a communication strategy characterized by the activity of artificial profiles on Twitter to amplify medical disinformation and push the reopening of activities in the United States. The investigation found that non-human Twitter accounts represented between 45% and 60% of the users who talk about covid-19. The study underlines that the level of bot involvement in other politicized events is normally between 10 and 20%. According to the study, many of these accounts were created in February 2020 and, during their lifespan, they have been used to disseminate and amplify disinformation, including false medical advice, conspiracy theories and propaganda to reopen commercial activities[113].

5. Ethical implications

As mentioned in the prior paragraphs, Artificial Intelligence applied to the healthcare sector raises a lot of ethical challenges that we'll address in the following sections. Among the various ethical issues arising from the use of Artificial Intelligence in this field, some are particularly intriguing and important to face. The European Parliament identified[114], in particular, two aspects: Patient privacy and confidentiality and universal access to equal quality of health care, but actually, the dimensions in which ethics emerged are more various.

5.1. Human rights and access

Artificial Intelligence can affect human rights in several different manners. One of the most important aspects is the universal access to Artificial Intelligence in medicine and equal and fair access to healthcare. In fact, residents of rural and remote locations, people living in developing countries as well as people on low-income may not have access to everything modern medicine has to offer since they may not have a well-equipped hospital close by, they cannot have enough resources to afford it or prices for health care and insurance may be prohibitively high. Moreover, vulnerable groups (e.g. minorities, disabled) may have less choice and lower levels of health care services.

The Council of Europe states[115] that even in exceptional circumstances, the quality of the services of the hospitals and healthcare structures should not be reduced nor compromise universal access to care and the support granted by information technology and Artificial Intelligence in healthcare shall not be seen only as a pathway to reduce cost. Indeed, it is necessary to consider the problems that could arise in terms of the quality of the services and global access to healthcare. Accordingly, Artificial Intelligence-driven healthcare should be guaranteed to everyone.

Coronavirus pandemic showed that even if AI systems helped in the COVID-19 fight, they are also able to negatively affect fundamental rights, as explained in the European Union's ethics guidelines for trustworthy AI. During the emergency, some countries took measures to control and mass

---

[112] B. Huang, K.M. Carley, Disinformation and Misinformation on Twitter during the Novel Coronavirus Outbreak, 2020, https://www.technologyreview.com/2020/05/21/1002105/covid-bot-twitter-accounts-push-to-reopen-america/, last visit January 2021

[113] The analysis has been performed by studying more than 200 million tweets related to coronavirus topic. The researchers used machine-learning and network analysis techniques to identify which accounts were spreading disinformation and which were most likely bots or cyborgs, accounts runned jointly by bots and humans.

[114] T. Evans, European framework on ethical aspects of artificial intelligence, robotics and related technologies, European parliament Research Service, 2020.

[115] Concil of Europe, AI and control of Covid-19 coronavirus, https://www.coe.int/en/web/artificial-intelligence/ai-and-control-of-covid-19-coronavirus, last visit January 2021.

monitoring the population by this technology and, moreover, AI was an instrument to spread fake news (see paragraphs 4.5 and 4.6). These are a few examples of how this new technology can directly affect human rights and, because of this, a fundamental rights impact assessment should be undertaken. "This should be done prior to the system's development and include an evaluation of whether those risks can be reduced or justified as necessary in a democratic society in order to respect the rights and freedoms of others. Moreover, mechanisms should be put into place to receive external feedback regarding AI systems that potentially infringe on fundamental rights"[116].

Moreover, standards relating to data protection, such as Convention 108(+) of the Council of Europe, should be fully applied under all circumstances like the use of biometric data, geolocalisation, facial recognition or the use of health data. Use of emergency measures should be carried out in full consultation with data protection authorities and respect the dignity and the private life of the users. The different biases of the various types of surveillance operations should be considered, as these may cause significant discrimination[117].

About the freedom of expression, the Council of Europe Committee of Experts on the Media Environment and Media Reform (MSI-REF) underlined in a statement[118] that "the crisis situation should not be used as a pretext to restrict public access to information. Nor should States introduce restrictions on media freedom beyond the limits allowed by Article 10 of the European Convention on Human Rights". The Committee also highlights that "member states, together with all media actors, should strive to ensure an environment conducive to quality journalism".

5.2. Confidentiality and informed consent

As already mentioned, AI applications for healthcare, such as imaging, diagnostics etc., are able to transform the patient-clinician relationship, and, in particular, they could affect the informed consent principle.

First of all, the confidential relationship between doctors and patients is essential for health care. Patient information should be available only to the patient, his/her treating physician and, to the degree required, other medical personnel (e.g. radiologist, nurse). This could be affected by the use of Artificial Intelligence, but it is not a problem if all the rules concerning privacy and security of the stored data will be respected. Problems may arise if patient data are shared among different AI in case these ones start to make a data treatment not understandable for humans. This happened, for example, during a Facebook experiment in which two AI appeared to be chatting to each other in a strange language only they understood[119].

Another aspect concerns the need to investigate under what circumstances the principles of informed consent should be applied when clinical procedures involve Artificial Intelligence. Particularly important is to determine to what extent do physicians have a responsibility to explain to the patient the AI complexities, their use in the procedures and, in particular, the types of ML used, the data used as input, the training dataset and the possibility of biases or shortcomings.

---

[116] High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

[117] A.F. Cahn, J. Veiszlemlein, COVID-19 tracking data and surveillance risks are more dangerous than their rewards, NBC News, 2020.

[118] Concil of Europe's Committee of experts on media environment and reform statement, 2020.

[119] The independent, https://www.independent.co.uk/life-style/facebook-artificial-intelligence-ai-chatbot-new-language-research-openai-google-a7869706.html, last visit January 2021.

About this, it has been valued that "patients are generally unaware if their physicians are using computerized decision aids to guide treatment. Indeed, patients receive little information about what sources their physicians consult"[120].

Find a solution to this issue is especially challenging in relation to cases in which AI operates using "black-box" algorithms. This kind of algorithm is "based on unknowable machine-learning techniques or because the relationships they draw are too complex for explicit understanding"[121]. Even if some of these techniques can help, others may be based on errors and lead to sub-optimal care for patients[122]. In this case, it is necessary to determine to what extent does a physician need to disclose that he cannot fully understand the processes above the diagnosis/treatment recommended by the AI and how much transparency is needed. Moreover, this "right to explanation"[123] has a connection with privacy (and in particular whit the EU's GDPR), since whether the patient may be reluctant to allow the use of certain categories of data (e.g., genetic data and family history), how it is possible to properly balance the privacy of patients with the safety and effectiveness of AI? Unfortunately, until now, there is no answer to this question, but it is clear that the two significant rights involved must be balanced. This, in turn, creates another unsolved problem. Who is the subject who has to decide?

Another issue related to informed consent concerns AI used in apps and chatbots. Such apps, indeed, raise serious questions about user agreements. Different from the traditional informed consent process, a user agreement is a contract agreed without a face-to-face dialogue[124].

As it happens for user agreements related to other kinds of apps, programmes and software, people (almost unanimously) do not spend the time to understand user agreements, often ignoring them. Moreover, continuous software updates make it difficult for individuals to understand adequately what terms of service they have agreed to[125].

The problem is then represented by the identification of the information that should be given to individuals using such apps, chatbots and software and by the methods that must be used to make the information understandable. Another side of the same problem is represented by users' education to carefully pay attention to this kind of issues. About this issue, the EU Parliament proposal about the creation of a license for AI users[126] could be interesting, but difficult to apply.

5.3. Safety and transparency

Safety represents one of the most important issues for AI in healthcare. As already explained Artificial Intelligence can provide output without being explicitly programmed. But what happens if the output is not correct?

---

120 I.G. Cohen, R. Amarasingham, A. Shah, B. Xie, B. Lo, The legal and ethical concerns that arise from using complex predictive analytics In health care, Health Affairs 2014, p. 1139-1147.

121 https://today.law.harvard.edu/petrie-flom-center-launches-project-precision-medicine-artificial-intelligence-law-pmail/, last visit January 2021.

122 For instance, Corti's algorithms are "black box" because even Corti's inventor does not know how the software reaches its decisions to alert emergency dispatchers that someone has a cardiac arrest.

123 A right to explanation (or right to an explanation) is a right to be given an explanation for an output of the algorithm. This right primarily refer to individual rights to be given an explanation for decisions that significantly affect an individual, including medical, legal or financial ones.

124 C. M. Klugman, L. B. Dunn, J. Schwartz, I. Glenn Cohen, 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, 2018, p. 38-47.

125 S. Gerke, T. Minssen, H. Yu, I. G. Cohen. Ethical and legal issues of ingestible electronic sensors. Nature Electron, 2019, p. 329-334.

126 European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), Ethical principles, n.11.

This is not a hypothetical situation, in fact, according to IBM internal documents[127], IBM Watson for Oncology[128], which uses AI algorithms to help physicians explore cancer treatment options for patients, recently gave "multiple examples of unsafe and incorrect treatment recommendations".

Probably the problem depends on the AI training phase. Instead of using real patient data, the software was only trained with a few synthetic cancer cases, meaning they were devised by doctors at the Memorial Sloan Kettering (MSK) Cancer Centre[129].

In order to avoid this kind of problems and to realize the full potential of AI, developers and other involved stakeholders have to ensure reliability and validity of the datasets on one hand and transparency on the other hand.

About the first issue, the datasets used for the training have to be reliable and valid, since better training data allow a better AI performance[130]. In addition, further refinement is necessary to generate more accurate results.

The second issue is represented by transparency. As a matter of fact, in the fields in which AI has to be extremely confident (such as in the medical sector, but, for example, also in driverless car one), high amounts of data and thus more data sharing are generally necessary, depending on the AI and the type of tasks it has to perform.

Moreover, despite the fact that there is a legitimate interest related to protecting investments and intellectual property and not increasing cybersecurity risk, some amount of transparency must be ensured. There are no rules, at this moment, to determine the transparency level, but a certain level of transparency must be guaranteed about the kind of data used and any shortcomings of the software (e.g., data bias). About this, the European Commission's High-level Expert Group on Artificial Intelligence in the Ethics guidelines for trustworthy AI indicates as requirements for transparency the following three elements: Traceability, Explainability and Communication[131].

The first one is related to the data sets and processes that yield the AI system's decision, which should be adequately documented and traceable to increase transparency. The same principle should be applied to the decisions made by AI systems. In this way, the identification of the reasons why a decision was erroneous could be easier and, consequently, could help prevent future mistakes. Explainability concerns the ability to explain the AI technical processes and related human decisions. This means that decisions taken by an AI system should be understood and traced by human beings.

Communication is related to the relationship between AI and human users. AI systems should not represent themselves as humans to users, which have the right to be informed that they are interacting with an AI system. Moreover, Artificial Intelligence capabilities and limitations should be communicated to all users.

In conclusion, transparency is able to create trust among all the involved stakeholders, and in particular, between clinicians and patients. This trust is a key element to permit a successful implementation of AI in clinical practice and this goal could be frustrated by the existence of "black-box" systems, as already explained (see paragraph 5.2).

---

127 https://www.beckershospitalreview.com/artificial-intelligence/ibm-s-watson-recommended-unsafe-and-incorrect-cancer-treatments-stat-report-finds.html and https://www.statnews.com/2018/07/25/ibm-watson-recommended-unsafe-incorrect-treatments/, last visit January 2021.

128 https://www.ibm.com/downloads/cas/0ZRYPWL9, last visit January 2021

129 S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, Artificial Intelligence in Healthcare, 2020, p. 295-336.

130 G.A. Kaissis, M.R. Makowski, D. Rückert et al, Secure, privacy-preserving and federated machine learning in medical imaging, Nature Machine Intelligence, 2020, 2, p. 305–311.

131 High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

5.4. Unfairness and bias

AI is believed to be a logical, neutral and impartial entity. Actually, any ML system or algorithm can be trustworthy and fair only if the data which it is trained with are[132]. According to this, AI can have biases and, consequently, it is able to discriminate. In particular, to avoid or, at least, minimize this risk, AI developers should consider this issue when deciding which kind of ML technologies/procedures they want to implement to train the algorithms and what datasets they intend to use, taking in consideration the quality and the diversity of the data.

Unfortunately, there are a lot of examples in which AI's biases provoked injustice with regard to ethnic origins, skin colour or gender[133,134,135]. Biases are not only related to race or gender, but they can also occur regarding other elements such as age or disabilities.

The existence of these biases is linked to multiple reasons, such as the unrepresentativeness of the datasets and the choice of methods and machine learning algorithms used by data scientists. Other sources of biases are represented by the context in which the AI is used[136], inadvertent historic bias and bad governance models. Moreover, AI developed in rich settings or developed countries could not recommend treatments that are accurate, safe, and fair in low-resource settings or in developing countries. Moreover, harm can also result from the intentional exploitation of (consumer) biases or by engaging in unfair competition[137] as a direct consequence of the will of the AI creator.

In the healthcare sector, if phenotype- or genotype-related information are involved, "biased AI could, for instance, lead to false diagnoses and render treatments ineffective for some subpopulations and thus jeopardize their safety"[138] due to the fact that AI is trained based on data focussed on Caucasian patients. In this case, the output could be less accurate or even inaccurate regarding other populations for which the training data was underinclusive.

This kind of biases can be avoided thanks to the increased data availability and, when this does not bring to an increased amount of data related to minorities, the specification for which populations the algorithm can or cannot be properly used can be useful.

But, also in this case, the main problem is represented by the "black-box" algorithms, in which the lack of transparency makes it impossible to detect biases. About this issue, it is interesting to underline that someone[139] sustains that the main issue is accuracy and when the decision is accurate it is not important that the processes behind the decision are transparent.

In order to avoid these biases, the European Union High-level expert group on AI suggests[140] the involvement, in the AI development, stakeholders who may directly or indirectly be affected by the

---

[132] High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

[133] A. Završnik, Criminal justice, artificial intelligence systems, and human rights, ERA Forum 20, 2020, p. 567–583.

[134] Price II, William Nicholson, Medical AI and Contextual Bias, 33 Harv. J.L. & Tech., University of Michigan Public Law Research Paper No. 632, 2019.

[135] O. Schwartz, In 2016, Microsoft's Racist Chatbot Revealed the Dangers of Online Conversation, IEEE spectrum, 2019, https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation, last visit January 2021.

[136] Price II, William Nicholson, Medical AI and Contextual Bias, 33 Harv. J.L. & Tech., University of Michigan Public Law Research Paper No. 632, 2019.

[137] High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

[138] S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, Artificial Intelligence in Healthcare, 2020, p. 295-336.

[139] A.J. London, Artificial Intelligence and Black-Box Medical Decisions: Accuracy versus Explainability, The Hasting Centre Reporting, 1, 2019, pp. 15-21.

[140] High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

system, looking for their feedback even after deployment of the algorithm. Moreover, the group advises to set up longer-term mechanisms for stakeholder participation, for example by ensuring workers information, consultation and participation throughout the whole process of implementing AI systems at organisations. This is good advice but it seems not enough to completely resolve the problem.

5.5. Privacy and data governance

Privacy is a fundamental right that could be particularly affected by AI applications. Consequently, it requires adequate data governance that can ensure that the quality of the data and protocols used is such as to protect privacy.

Consequently, Artificial Intelligence systems should guarantee privacy and data protection during their whole lifespan. According to the EU High-level expert group on AI "this includes the information initially provided by the user, as well as the information generated about the user over the course of their interaction with the system (e.g. outputs that the AI system generated for specific users or how users responded to particular recommendations). Digital records of human behaviour may allow AI systems to infer not only individuals' preferences but also their sexual orientation, age, gender, religious or political views. To allow individuals to trust the data gathering process, it must be ensured that data collected about them will not be used to unlawfully or unfairly discriminate against them"[141].

Moreover, any organisation that handles individuals' data should create protocols governing data access. "These protocols should outline who can access data and under which circumstances. Only duly qualified personnel with the competence and need to access individual's data should be allowed to do so"[142].

Privacy in the healthcare system is related to health data, which value can reach up to billions of dollars in a market in which companies sell patient data for profit. This means that every user must be informed about which kind of data is taken when they use a particular program, app, AI, and he must give the related informed consent (see paragraph 5.2).

It is also very important to protect patients from data usage outside the doctor-patient relationship. Data gathered for medical treatment can affect directly patients in several aspects of their life, like insurance issues, job opportunities, or even personal relationships[143]. Additionally, some AI health apps also raise further issues, since health data can be shared not only with the doctor but also with family members or friends (apple watch[144], family sharing[145]), that are not subject to duties of confidentiality.

Lastly, another issue is related to automated processes. Article 22 of the GDPR limits the circumstances in which it is possible to make solely automated decisions able to have a legal or similarly significant effect on individuals. Solely means a decision-making process that is totally automated and excludes any human influence on the outcome (see paragraph 6.3).

Accordingly, AI technologies in health care are directly concerned by this provision that implies that a system created to provide as output a treatment suggestion cannot be used as the sole basis for the prescription.

---

141 High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

142 High-level expert group on Artificial Intelligence, Ethics guidelines for trustworthy AI, Requirements for trustworthy AI, 2019.

143 M. B. Forcier, H. Gallois, S. Mullan, Y. Joly, Integrating artificial intelligence into health care through data access: can the GDPR act as a beacon for policymakers?, Journal of Law and Bioscience, 2019, 6, p. 317-335.

144 https://support.apple.com/en-us/HT207014, last visit January 2021.

145 https://www.apple.com/family-sharing/, last visit January 2021.

6. Juridical issues

All the above-discussed issues present legal implications, but the regulation of Artificial Intelligence is still in a de iure condendo stage and several legislators, national and supranational, are still moving the first steps.

Main legal issue/area	applicable EU legislation
Algorithmic transparency	Regulation 2016/679; Directive 206/680
Unfairness, bias and discrimination	Article 2, 3(3), 9 TEU; Article 10 TFEU; Article 20-26 EU Charter on Fundamental Rights; Council Directive 200/78/EC; Council Directive 2000/43/EC; Council Directive 2004/113/EC; Directive 2006/54/EC, Council Directive 79/7/EEC, Directive 2010/41/EU, Council Directive 2010/18/EU, Regulation (EU) 2016/679, Directive (EU) 2016/680, Directive (EU) 2016/681.
Intellectual property issues	Article 118 TFEU, Article 17 (2) EU Charter on Fundamental Rights; Directive 2001/29/EC; Directive 2006/115/EC; Directive 2001/84/EC; Directive 2009/24/EC; Directive 2004/48/EC; Directive 96/9/EC; Directive 2012/28/EU; Directive 98/71/EC; Regulation (EU) No 1257/2012; Regulation (EU) 2017/1001; Directive (EU) 2016/943;
Legal personhood of AI	Not covered[146].
Vulnerability and cybersecurity	Directive (EU) 2016/1148; Regulation (EU) No 910/2014; Directive 2013/40/EU; Regulation (EU) No 526/2013; Directive 2002/58/EC
Impact of AI on workers	Article 3(1)(3) TEU; Article 9, 107(3)(a), Articles 145-166 TFEU; Articles 14-15, 27-32 EU Charter of Fundamental Rights; Regulation (EU) No 1304/2013
Privacy and data protection	Articles 7-8 EU Charter of Fundamental Rights; Regulation (EU) 2016/679; Directive (EU) 2016/680; Directive (EU) 2016/681; Directive 2002/58/EC
Liability	Articles 4(2)(f), 12, 114 and 169 TFEU; Articles 38, 47 EU Charter of Fundamental Rights; Council Directive 85/374/EEC
Accountability for harm	Regulation (EU) 2016/679

Table 1: Applicable EU legislation to legal issue related to AI

In the following paragraphs, it will be presented an overview of the European Union approach to some of the main legal issues related to Artificial Intelligence application in the medical field, since AI may either enhance or impair the exercise of multiple fundamental rights: privacy and data protection, civil freedoms and social rights[147].

6.1. Civil liability

[146] About this topic, see Daniele Chiappini, Artificial Intelligence and legal capacity: an introduction to robotic subjectivity, Diritto e Processo, 2019.
[147] EPRS (European Parliament Research Service), The impact of General Data Protection Regulation (GDPR) on Artificial Intelligence, 2020.

The existing legal framework on civil liability is applicable to all advanced technologies and it is primarily based on fault (or culpa aquiliana) and contractual liability rules.

"Fault-based liability establishes the duty to compensate damages upon the subject who negligently failed to maintain a desired standard of behaviour, causing damages. Vicarious liability and other forms of indirect liability rules – for things or animals in custody, or for the acts of children and auxiliaries – could be applied, holding primarily owners, and users of advanced technologies liable.

Contractual liability, instead, presupposes a legally qualified relationship (contractual or legal obligation) between the parties before damage occurs"[148].

Despite these common elements, new AI-based technologies raise challenges for current liability regimes that are differently addressed in different legislations. In particular, concerning the medical sector in the EU, one of the main aspects that need to be analysed is how liability might be apportioned between the physicians, the manufacturer, the programmer and the hospital or structure adopting the AI-based tool, and/or employs the practitioner. To this purpose, an extensive body of existing EU product safety and liability legislation, including sector-specific rules, further complemented by national legislation, is relevant and potentially applicable to a number of emerging AI applications[149].

The European Parliament stated that the more autonomous robots are, the less they can be considered to be simple tools in the hands of other actors[150]. This, in turn, poses a question of whether the existent rules on liability are sufficient or whether it is necessary to create new principles and rules to provide clarity on the legal liability of the previously mentioned actors. The main problem concerns, mainly, the responsibility for the acts and omissions of robots where the cause cannot be traced back to a specific human actor and the responsibility related to situations in which the acts or omissions of robots that have caused harm could have been avoided.

Unfortunately, an EU study[151] clarified that no one-size-fits-all approach may be adopted, even within a relatively narrow-tailored class of AI applications. For example, under medical robots or medical AI sectors fall a broad variety of devices, which are quite different for technical features, diffusion, function and use.

In the current system, the European Commission believes[152] that, in a future regulatory framework, each obligation should be addressed to the actors who are best placed to address any potential risks, since a lack of clear safety provisions tackling these risks may create legal uncertainty.

In the Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics[153], which accompanies the Commission's White Paper, it is stated that provisions explicitly covering risks caused by the emerging digital technologies could be introduced since technologies like AI may challenge some aspects of the liability frameworks and could reduce their effectiveness. Moreover, some of these characteristics could make it hard to trace the damage back to a person, which is necessary for a fault-based liability in accordance with most national rules. This could

---

148 A. Bartolini, Artificial Intelligence and Civil Liability, JURI Committee study, Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, PE 621.926, 2020, p. 10.

149 European Commission, White paper on Artificial Intelligence – A European approach on excellence and trust, 2020.

150 European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), Liability, letter AB.

151 A. Bartolini, Artificial Intelligence and Civil Liability, JURI Committee study, Policy Department for Citizens' Rights and Constitutional Affairs Directorate-General for Internal Policies, PE 621.926, 2020, p. 111.

152 European Commission, White paper on Artificial Intelligence – A European approach on excellence and trust, 2020, p. 22.

153 European Commission, Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, 2020.

significantly increase the costs for victims and meant that liability claims against others than producers may be difficult to make or prove.

Consequently, according to the European Commission's intent, the new regulatory framework for AI should be able to achieve its objectives without being excessively prescriptive, following a risk-based approach. This kind of approach requires setting clear criteria to understand which Artificial Intelligence applications are included in the high-risk group and which are not. AI application not qualified as high-risk will remain entirely subject to already existing EU-rules.

To respond to this main issue, the Commission proposes that a given AI applications should generally be considered high-risk "in light of what is at stake, considering whether both the sector and the intended use involve significant risks, in particular from the viewpoint of protection of safety, consumer rights and fundamental rights"[154] meeting the following two cumulative criteria:

- The sector in which Artificial Intelligence is applied, given the characteristics of the activities typically undertaken, is a sector in which it is likely that significant risks can occur. These sectors should be specifically and exhaustively listed in the new regulatory framework, and healthcare is one of the sectors used as an example among the high-risk ones.
- The AI applications in the listed sector are used in such a manner that significant risks are likely to arise. This because not every use of AI in the selected sectors necessarily involves significant risks. The white paper specifies, for the medical sector, that healthcare generally is a relevant sector, but a flaw in the appointment scheduling system in a hospital will normally not pose risks of such significance as to justify legislative intervention. The assessment of the level of risk of a given use could be based on the impact on the affected parties, for instance, in medical sector, each use of AI applications that pose a risk of injury or death.

The European parliament's IMCO Committee[155] underlined in its opinion on the Civil liability regime of artificial Intelligence[156] the need for a risk-based approach within the existing liability framework, sharing the view of the European Commission. It is interesting that the Committee asks the Commission to assess the introduction of a separate yet complementary strict liability regime for AI systems presenting a high risk of causing harm or damage to a person or property in a manner that is random and impossible to predict in advance. This regime should ensure that victims are effectively compensated for damage caused by AI-driven systems[157].

Anyway, it is important guarantee that, independently of the rules chosen, the new regulatory system shall aim to assure that people harmed by this new technology receive the same level of protection compared to victims of traditional technologies.

6.2. AI Implementation in Medical Devices

In paragraph 5.2, it has been shown the importance of a safe and effective implementation of Artificial Intelligence-based medical devices. This implementation passes through regulation of the

---

[154] European Commission, Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee, Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics, 2020, p.17.

[155] The IMCO Committee is the European Parliament Committee on the Internal Market and Consumer Protection.

[156] European Parliament Committee on the Internal Market and Consumer Protection, Opinion of the European Parliament Committee on the Internal Market and Consumer Protection for the Committee on Legal Affair on Civil liability regime for Artificial Intelligence of 7 July 2020.

[157] European Parliament Committee on the Internal Market and Consumer Protection, Opinion of the European Parliament Committee on the Internal Market and Consumer Protection on Civil liability regime for Artificial Intelligence, 2020, art. 20.

sector by the law. Despite the fact that it does not exist specific regulations devoted to AI, current European rules on medical devices are applicable.

In particular, it is pivotal the Medical Device Regulation[158] (MDR). It was supposed to become effective on May 26, 2020, but, due to the pandemic COVID-19, it will enter into force in 2021. It will repeal the Medical Device Directive (93/42/EEC – MDD) and the Directive on active implantable medical devices (90/385/EEC – AIMD)[159].

The new MDR makes some changes to the classification of medical devices, including in medical device definition some software that, under the previous directive, was not included. Accordingly, the new medical device definition in Art. 2 (1) comprehends software used for the Medical purpose of prediction or prognosis of disease as a medical device. Anyway "software for general purposes, even when used in a healthcare setting, or software intended for lifestyle and well-being purposes is not a medical device"[160].

Similar to the previous Directive, MDR classifies medical devices into four categories[161], and in particular, classes I, II a, II b, and III, based on the intended purpose of the medical devices and their inherent risks. The regulation introduces new implementing and classification rules explicitly focussed on software. According to this rule, "software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class II a, except if such decisions have an impact that may cause:

- death or an irreversible deterioration of a person's state of health, in which case it is in class III or
- a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb.

Software intended to monitor physiological processes is classified as class IIa, except if it is intended for monitoring vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as class IIb.

All other software is classified as class I"[162].

These new rules entail a reclassification of software and, accordingly, in October 2019, the Medical Device Coordination Group released nonbinding guidance on qualification and classification of software under the regulation[163].

The new Regulation will be accompanied by a CE marking attesting the conformity of the device with the requirements set out in the MDR so that it can move freely within the EU[164]. In particular, manufacturers of medical devices shall undertake an assessment of the conformity of their devices prior to placing them on the market[165]. This evaluation must be based on the classification and type of the particular device[166]. For example[167], class I devices have a low level of vulnerability and thus the

---

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

159 Art. 122 of the Medical Devices Regulation (2017/745).

160 Regulation (EU) 2017/745 on Medical Devices, Recital 19.

161 Regulation (EU) 2017/745 on Medical Devices, art. 51 (1).

162 Regulation (EU) 2017/745 on Medical Devices, Annex VIII, Chapter III, Rule 11.

163 Medical Device coordination group document, MDCG 2019-11, Guidance on Qualification and Classification of Software in Regulation (EU) 2017/745 – MDR and Regulation (EU) 2017/746 – IVDR, 2019.

164 Regulation (EU) 2017/745 on Medical Devices, Recital 40, Regulation (EU) 2017/745 on Medical Devices, Art.2 (43).

165 Regulation (EU) 2017/745 on Medical Devices, Art. 52, Regulation (EU) 2017/745 on Medical Devices, Annexes IX – XI.

166 Regulation (EU) 2017/745 on Medical Devices, Art. 52.

conformity assessment procedure can generally be carried out under the sole responsibility of the manufacturers[168]. In contrast, class II a, II b, and III devices that have a higher risk than class I devices entail the involvement of a notified body, a conformity assessment body designated in accordance with the MDR[169].

6.3. Privacy and data protection

The General Data Protection Regulation (GDPR) entered in force in 2018 affirming the right of natural persons to the protection of personal data[170]. It is applied to the "processing of personal data in the context of the activities of an establishment of a controller or a processor" in the EU, notwithstanding of whether the processing takes place in an EU or non-EU country[171]. Moreover, the Regulation is applied also in cases where the processor or controller is established in a non-EU country and processes "personal data of data subjects who are in the Union" for "the offering of goods or services" to such data subjects in the EU or for the "monitoring" of the data subjects' behaviour[172].

The GDPR defines "personal data" as "any information relating to an identified or identifiable natural person ('data subject')"[173] and "processing" as "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means," including collection, structuring, storage, or use[174]. "Controller" is "the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data," while "processor" is "a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller"[175].

In the healthcare context, GDPR defines "data concerning health"[176] as "personal data related to the physical or mental health of a natural person, including the provision of healthcare services, which reveal information about his or her health status." The EU's GDPR is a lot broader in its scope compared to the US' HIPAA[177], which only covers specific health information generated by "covered entities" or their "business associates"[178].

Moreover, the processing of special categories of personal data such as genetic data, biometric data, and data concerning health is generally prohibited[179], but there are explicit exceptions[180] such as the cases, relevant for the healthcare sector, in which "the data subject has given explicit consent […] for one or more specified purposes", where the "processing is necessary for reasons of public interest

---

167 S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, Artificial Intelligence in Healthcare, 2020, p. 295-336.
168 Regulation (EU) 2017/745 on Medical Devices, Recital 60, Regulation (EU) 2017/745 on Medical Devices, art. 52 (7).
169 Regulation (EU) 2017/745 on Medical Devices, Recital 60, Regulation (EU) 2017/745 on Medical Devices, art. 2 (42).
170 Regulation (EU) 2016/679 on General Data Protection, art. 1 (2).
171 Regulation (EU) 2016/679 on General Data Protection, arts. 2,3 (1).
172 Regulation (EU) 2016/679 on General Data Protection, art. 3 (2).
173 Regulation (EU) 2016/679 on General Data Protection, art. 4 (1).
174 Regulation (EU) 2016/679 on General Data Protection, art. 4 (2).
175 Regulation (EU) 2016/679 on General Data Protection, art. 4 (7) and (8).
176 Regulation (EU) 2016/679 on General Data Protection, art. 4 (15).
177 United States Health Insurance Portability and Accountability Act. It was created in 1996 to modernize the flow of healthcare information, stipulate how personally identifiable information maintained by the healthcare and healthcare insurance industries should be protected from fraud and theft, and address limitations on healthcare insurance coverage.
178 S. Gerke, T. Minssen, G. Cohen, Ethical and legal challenges of artificial intelligence-driven healthcare, Artificial Intelligence in Healthcare, 2020, p. 295-336.
179 Regulation (EU) 2016/679 on General Data Protection, art. 9 (1).
180 Regulation (EU) 2016/679 on General Data Protection, art. 9 (2).

in the area of public health" or "for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes"[181].

GDPR provisions are applicable to AI-driven medicine since, in the collection of personal data, the controllers must provide data subjects with information about "the existence of automated decision-making, including profiling, referred to in Article 22 (1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"[182], where for "automated decision-making" is intended the ability to make decisions by technological means without human involvement[183].

Moreover, the right of access to the personal data recognised to data subject concerns also the information about "the existence of automated decision-making, including profiling [and] meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"[184].

There are two kinds of information, ex-ante, addressed by the right to information established by Articles 13 (2) letter f and 14 (2) letter g and ex-post, addressed by Article 15 (1). The former establishes that must be provided two kinds of information, one on the existence of automated decision-making and one on its logic and envisaged consequences. About this, it is useful to underline that there is a conflict between the need for the information to be concise and understandable on the one hand and the need for it to be precise and in-depth on the other.

The latter (ex-post information) reiterates the same information requirements in Articles 13 and 14, but it must be determined whether the controller is required to provide the data subject with only general information or also with an individualised explanation[185].

Particularly important is the right recognised to data subjects to "not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."[186] Exceptions to what established in art. 22 (1), provided in Article 22 (2) of the GDPR, are not applicable for decisions based on genetic and biometric data as well as data concerning health[187]. This means that it is not possible to take a decision only basing on the output of an AI elaboration, but it is necessary the intervention of a physician who shall value the AI suggestion and take responsibility for the adopted procedure.

One of the most interesting and controversial issues is related to the right to explanation and its connection with the black box problem. Recital 71 of the GDPR recognises the right "to obtain an explanation of the decision reached after such assessment and to challenge the decision." The main problem is related to the legal existence of such a right to an explanation of specific automated decisions since Recital 71 of the GDPR is not legally binding. This is deeply connected with the already explained "black box" problem (see paragraph 5.2). If such a right would be recognised, data subjects would be entitled to obtain "meaningful information about the logic involved, as well as the significance and the envisaged consequences" of automated decision-making systems, with the consequent necessity to open the above-mentioned black box.

---

181 Regulation (EU) 2016/679 on General Data Protection, art. 9 (2), letters a, I and j.

182 Regulation (EU) 2016/679 on General Data Protection, art. 13 (2) letter f and art. 14 (2) letter g.

183 Working Party on the Protection of Individuals with Regard to the Processing of Personal Data, Guidelines on automated individual decision-making and profiling for the purposes of Regulation 2016/679, wp251rev01, p. 8.

184 Regulation (EU) 2016/679 on General Data Protection, art. 5 (1) letter h.

185 EPRS (European Parliament Research Service), The impact of General Data Protection Regulation (GDPR) on Artificial Intelligence, 2020, par 4.2.4 AI and transparency, p. 74.

186 Regulation (EU) 2016/679 on General Data Protection, art. 22 (1).

187 Regulation (EU) 2016/679 on General Data Protection, art. 22 (4).

Anyway, eminent tenet[188] sustains that this right cannot be recognised, due to the absence of a legal binding of recital 71, and we can agree with this position.

7. – Conclusions

In this paper, we have given an overview of the state of the art on what AI is, how it is already applied in the healthcare sector and about some of the ethical and legal issues related to AI in the healthcare sector.

In particular, the main issues discussed examine some of the primary ethical challenges that need to be addressed, such as human rights and access to AI, confidentiality and informed consent, safety and transparency, unfairness and biases and privacy and data governance. This has been followed by an analysis of the main legal challenges and in particular civil liability, implementation, and privacy and data protection.

According to the overview, there are a lot of elements that are key factors necessary to create a reliable and successfully AI-based healthcare system, namely, informed consent, high levels of data protection and privacy, algorithmic fairness, transparency and regulatory oversight, high standards of safety and an optimal liability regime for AI.

The impact of Artificial Intelligence in everyday life is so deep that it is necessary to rethink current regulatory frameworks and update them to the new technological developments, as the European Union is doing. As a matter of fact, European institutions, such as the European Parliament and the European Commission are currently working on new regulatory frameworks related to Artificial Intelligence in general and also to singular aspects of AI.

Moreover, it is very important to have public and political discussions centred on the ethics of AI-driven healthcare such as its implications on the human workforce and the society since AI has a high potential to improve the healthcare system. To do so it is necessary not only to address the arising ethical and legal challenges but also to form the users about the implication of AI usage to avoid that a lack of knowledge leads to fear and opposition to artificial intelligence use. This can be adequately realized only if the legislators, national and supranational, will be able to face the issues in time and with the correct perspective, a perspective respectful of human rights and able to find a solution to possible problems arising from the AI applications.

---

[188] S. Wachter, B. Mittelstadt, L. Floridi, Why a right to explanation of automated decision-making does not exist in the general data protection regulation, International Data Private Law, 2017.