THE DEPLOYMENT OF ARTIFICIAL INTELLIGENCE TOOLS IN THE HEALTH SECTOR: PRIVACY CONCERNS AND REGULATORY ANSWERS WITHIN THE GDPR

Mirko Forti*

This article examines the privacy and data protection implications of the deployment of machine learning algorithms in the medical sector. Researchers and physicians are developing advanced algorithms to forecast possible developments of illnesses or disease statuses, basing their analysis on the processing of a wide range of data sets. Predictive medicine aims to maximize the effectiveness of disease treatment by taking into account individual variability in genes, environment, and lifestyle. These kinds of predictions could eventually anticipate a patient's possible health conditions years, and potentially decades, into the future and become a vital instrument in the future development of diagnostic medicine. However, the current European data protection legal framework may be incompatible with inherent features of artificial intelligence algorithms and their constant need for data and information. This article proposes possible new approaches and normative solutions to this dilemma.

Keywords: artificial intelligence, algorithm, medicine, health, privacy, data protection, GDPR

I. INTRODUCTION

The deployment of Artificial Intelligence (AI) instruments in the medical sector has led to significant diagnostic innovations. AI tools, due to their capacity to elaborate vast amonts of data in real-time, can individuate common patterns undetectable for human physicians.[1] The global diffusion of mobile and wearable technology, like smartphones and smartwatches, enables the collection and uploading of vast amounts of data into AI

* Research Fellow - Sant'Anna School of Advanced Studies – mirko.forti@santannapisa.it

1 Charles A.Taylor and others, 'Predictive Medicine: Computational Techniques in Therapeutic Decision-Making' (1999) 4 Computer Aided Surgery 231.

algorithms.[2] These technological instruments take into account an extensive array of patients' genetic features to provide tailored medical treatments. A relatively new speciality of the medical field called predictive medicine[3] involves the processing of genetic and laboratory tests using AI tools to predict the outbreak of a disease. Thus, the control and ownership of data are particularly relevant legal concerns for medical care.

The General Data Protection Regulation[4] (GDPR, the Regulation) is the core of the European Union's (EU) approach regarding privacy and data protection. It formulates normative standards with which algorithms must also comply. However, the GDPR dates back to an era when AI algorithms did not yet play a fundamental role in everyday life. As a result, central components of AI tools may raise compliance concerns. Most significantly, the reasoning routine of AI algorithms is obscure and undetectable due to the inherent opaqueness of AI tools. An external human observer cannot detect and recreate the reasoning pattern chosen by the algorithm system, even if the output delivered by the AI tool is available.[5] This lack of understanding and reproducibility, known as the 'black-box' status of AI,[6] is incompatible with the fundamental requirements of transparency, fairness and accountability enshrined in the GDPR to ensure lawful and legitimate processing of personal data.

However, opening the black box would mean showing the functioning mechanism of the algorithm to market competitors, which could stifle

---

2 Jason P.Burnham and others, 'Using Wearable Technology to Predict Health Outcomes: a Literature Review (2018) 25 Journal of the American Medical Informatics Association 1221.

3 Maxwell Y. Jen and others, 'Predictive Medicine' [2020] StatPearls <https://www.ncbi.nlm.nih.gov/books/NBK441941/> accessed 15 July 2020.

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

5 Robin C. Feldman et al.,and others, 'Artificial Intelligence in the Health Care Space: How We Can Trust What We Cannot Know', (2019),),) 30, Stanford Law and Policy Review, 399.

6 Yavar Bathaee, 'The Artificial Intelligence Black Box and the Failure of Intent and Causation' (2018) 31 Harvard Journal of Law and Technology 889.

innovation unless more transparent forms of AI became eligible for patent protection.[7] Moreover, making the working processes of this software detectable would be technically impossible: AI algorithms continuously change their working routine to follow new patterns and perform tasks in an ever more efficient way.[8] Changing their computing patterns allows them to produce more reliable diagnostic outcomes. Accordingly, a human observer would not be able to understand the reasoning process of AI tools and its relationship to the data processed, even if these were visible. Ultimately, this lack of interpretability[9] can result in a lack of trust in the effective functioning of AI algorithms because researchers and physicians must base their clinical decisions on the correct functioning of black-box algorithms.

Can the current European legal framework adequately address the main privacy-related issues that arise from the use of AI software for diagnostic purposes? More specifically, can the European Regulation foster the development of predictive medicine and, at the same time, protect the rights of patients involved in the medical treatments? Starting from previous scholarship, such as the work of Ann Cavoukian regarding the principle of privacy by design,[10] this article seeks to find normative solutions within the GDPR to address the deployment of AI tools in the medical sector. Furthermore, it attempts to find a point of balance between two opposing interests: on the one hand, the privacy rights of every individual and, on the

---

7 See Ana Ramalho, 'Patentability of AI-Generated Inventions: is a Reform of the Patent System Needed?' (2018) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3168703> accessed 19 November 2020.

8 Jenna Burrell, 'How the Machine "Thinks": Understanding Opacity in Machine Learning Algorithms' (2016) 1 Big Data and Society 10.

9 Feldman (n 5); William J. Murdoch and others, 'Definitions, Methods and Applications in Interpretable Machine Learning' (2019) 116 PNAS 22071.

10 Ann Cavoukian, 'Privacy by design: the 7 foundational principles. Implementation and mapping of fair information practices' (*Data Protection Industries*) <http://dataprotection.industries/wp-content/uploads/2017/10/privacy-by-design.pdf> accessed 24 February 2020.

other hand, the general interest to stimulate scientific and medical research and, in more general terms, the right to public health.[11]

The article begins by addressing the most relevant norms of the GDPR to highlight the privacy-related issues arising from the deployment of AI tools in the medical sector. More specifically, it focuses on the main legal uncertainties arising from incompatibilities between the use of AI tools for predictive medicine and norms like the principles of transparency, fairness and lawfulness, the issue of free, informed and specific consent, the right to be forgotten, and the prohibition of automated decisions. The article then provides a few reflections about the main privacy threats raised by the development of predictive medicine and how to overcome them. The inherent opaqueness of AI algorithms may present a challenge for the transparent and lawful functioning of predictive medicine. However, the concepts of privacy by design and privacy by default[12] could represent the normative basis on which AI algorithms can be made compliant with the provisions of the GDPR.

II. The Principles of Fairness, Transparency and Lawfulness Applied to Data Processing in the Medical Field

The recent approval and entry into force of the GDPR established new privacy standards for data protection at a European level.[13] While in one respect the GDPR actually reduces obligations for data controllers regarding access to clinical data compared to previous legislation, it also limits utilisation of health data without consent, regulates its secondary use (that is,

---

11 Shane O'Sullivan and others, 'Legal, Regulatory and Ethical Frameworks for Development of Standards in Artificial Intelligence (AI) and Autonomous Robotic Surgery' (2019), 15(1) The International Journal of Medical Robotics and Computer Assisted Surgery 1.

12 GDPR, art 25.

13 Mélanie Bourassa Fourcier and others, 'Integrating artificial intelligence into health care through data access: can the GDPR act as a beacon for policymakers?' [2019] Journal of Law and the Biosciences 317.

use of data for purposes beyond those originally planned),[14] and requires data processing activities in all fields, including predictive medicine, to comply with certain fundamental principles. The principles of fairness, transparency and lawfulness are the cornerstones of the current European privacy legal framework.[15] They form the main thread uniting all processing activities and ensure the protection of the fundamental rights of people involved.

1. The Principle of Fairness in the Functioning of AI Algorithms

The principle of fairness is central to the relationship between the controller and the data subject[16] and is particularly crucial in the functioning of AI algorithms. These predictive tools may exacerbate discriminatory trends if they process prejudicial data. In the healthcare sector this may have fatal consequences for patients. Discriminatory factors, such as race[17]or gender,[18] could shape the final predictive outcome, implicating the ethical obligation of non-maleficence[19] according to which every medical treatment should promote patient safety and recovery.[20] Meanwhile, relying on a neutrality[21] conception for AI algorithms, where these tools produce standardized (neutral) outcomes ignoring the peculiar differences between patients, could

---

14 William Lowrance, 'Learning from the Experience: Privacy and the Secondary Use of Data in Health Research', (2003), 8 Journal of Health Services Research & Policy 2.

15 GDPR, art 5.1

16 European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European Data Protection Law* (Publications Office of the European Union 2018) 118.

17 Ziad Obermeyer and others, 'Dissecting racial bias in an algorithm used to manage the health of populations' (2019) 366 Science 447.

18 Davide Cirillo and others, 'Sex and Gender Differences and Biases in Artificial Intelligence for Biomedicine and Healthcare' (2020) 3(81) NPJ Digital Medicine 1.

19 Gunnar B.J. Andersson and others, 'Do Not Harm: The Balance of "Beneficence" and "Non-Maleficence"' (2010) 35(9S) Spine S2; Vittorio Bufacchi, 'Justice as Non-Maleficence' (2020) 67(162) Theoria 1.

20 Melissa D.McRadden and others, 'Ethical Limitations of Algorithmic Fairness Solutions in Health Care Machine Learning, (2020) 2 The Lancet Digital Health E221.

21 Ruha Benjamin, 'Assessing Risk, Automating Racism' (2019) 366(6464) Science 421.

be likewise detrimental.[22] The challenge is to design AI algorithms capable of taking into account environmental and societal factors[23] and inherent biological differences to provide fair and reliable health predictive outcomes. Not all subjects react in the same way as the average model to a specific medical treatment. In other words, fairness does not mean equality at all costs. A "fair" algorithm deployed for diagnostic purposes should be aware of the limitations of model predictions caused by social determinants of health and biological peculiarities.

Overcoming the issue of biased outputs may require human intervention. Physicians can reformulate medical questions to reduce bias.[24] They can rely on causal knowledge to verify the algorithmic decision and identify medical problems where the consequences of datasets biases are relatively negligible; in other words, reformulating the input to generate a fairer output. Developing guidelines to standardize reporting of predictive models delivered by AI algorithms can also reduce discrepancies between outcomes and help validate diagnostic products.[25] Thus, physicians can choose in a transparent way which kind of algorithms they should use according to the peculiarities of the specific medical case.

2. The Principle of Transparency as a Safeguard for the Privacy Rights of Data Subjects

The principle of transparency requires the controller to inform the data subject about every phase of the processing operation and explain these phases in a clear and understandable way. Transparency in data processing is strictly correlated with the principle of purpose limitation. In order for a data subject to exercise their privacy rights, the individual must know the reasons

---

22 McRadden (n 20).

23 Michael Marmot, 'Social Determinants of Health Inequalities' (2005) 365(9464) The Lancet P1099.

24 Nanette K.Wenger, 'Cardiovascular Disease: The Female Heart Is Vulnerable. A Call to Action grom the 10Q Report' (2012) 35 Clinical Cardiology 134.

25 Gary S.Collins and others, 'Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement' (2015) 13 BMC Medicine <https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-014-0241-z> accessed 23 November 2020.

for which their data is being collected and processed.[26] Processing without a specific, determined goal is unlawful. However, due to the black-box nature of AI algorithms, scientists and physicians can neither *ex ante* inform their patients regarding every possible outcome of the AI working process nor forecast possible future uses of data already elaborated.

In the medical sector, the goal is to make the diagnostic routine more user-centric, protecting the identity rights of the patients. New technological approaches could help.[27] For instance, so-called federated learning (a machine learning technique to process data through decentralized devices, in which each server assembles its own dataset) facilitates collaborations across multiple institutions without sharing patient data.[28] Data controllers can adjust and improve the effectiveness of their data processing model and then share these improved algorithms with other subjects through a trusted server, thereby obtaining a trained AI algorithm without sharing personal data with third parties. Similarly, the advanced technique of homomorphic encryption allows algorithms to elaborate data without decoding encrypted information,[29] and thus without identifying the underlying data subject.

3. The Principle of Lawfulness and the Secondary Use of Data

The principle of lawfulness requires each data processing procedure to be grounded on one of six legal bases specified in the GDPR. These include, *inter alia*, performing a contract, protecting the vital interest of a person, or

---

26 Article 29 Working Party, 'Opinion 03/2013 on Purpose Limitation' (2 March 2013) 00569/13/EN WP 203 <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 28 February 2020.

27 Casimir Wierzynski, 'Advancing Both AI and Privacy Is Not a Zero-Sum Game' (*Fortune*, 27 December 2018) <https://fortune.com/2018/12/27/ai-privacy-innovation-machine-learning/> accessed 21 July 2020.

28 Micah J.Seller et al., 'Federated Learning in Medicine: Facilitating Multi-Institutional Collaborations without Sharing Patient Data' (2020) 10 Scientific Reports <https://www.nature.com/articles/s41598-020-69250-1> accessed 19 November 2020.

29 Mohamed Alloghani et al., 'A systematic review on the status and progress of homomorphic encryption technologies', (2019),48, Journal of Information Security and Applications, 1.

safeguarding public interests. The secondary use of sensitive data is considered lawful when the processing activities are for scientific purposes,[30] including reasons of public health, but only if the utilised dataset does not permit identification of any data subject previously involved.[31] Using the same dataset for several purposes is fundamental to the correct development of scientific research,[32] but the strict normative provisions of the GDPR could discourage scientists and physicians from fully exploiting the research possibilities intended by the European legal framework.[33]

III. Freely-given, Specific and Informed Consent in Black-box Medicine

Consent is central to data-elaborating activities. The collection of so-called sensitive data, such as information regarding an individual's health conditions, is prohibited unless there is room to apply one of the exemptions listed by the GDPR, including explicit consent.[34] Such consent requires the free, informed, specific and unambiguous indication of the agreement stated by data subjects regarding the processing of their personal data.[35] Data subjects must have a real choice to provide legitimate consent,[36] and thus must be aware of specific details regarding the processing activities. These include the identity of the data controller, the purposes of every operation for which they gave their consent, the possibility to withdraw consent at any time, without experiencing technical difficulties,[37] and information about the use of their data for automated decision making, if applicable.[38] Furthermore,

---

30 GDPR, art 89.1 of the GDPR.

31 GDPR, recital 156.

32 Gauthier Chassang, 'The Impact of the EU General Data Protection Regulation on Scientific Research' (2017) 11 ecancermedicalscience 709.

33 Fourcier (n 13).

34 GDPR, art 9.

35 GDPR, art.4.11; Mary Donnelly, Maeve McDonagh, 'Health Research, Consent and the GDPR Exemption' (2019) 26 European Journal of Health Law 97.

36 European Data Protection Board, 'Guidelines 05/2020 on Consent under Regulation 2016/679' (4 May 2020) <https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202005_consent_en.pdf> accessed 20 November 2020.

37 GDPR art 7.

38 European Data Protection Board (n 36).

the data subject should be able to understand every feature and characteristic of the processing procedures.[39] However, the GDPR does not state anything in terms of competence and capacity of the data subject.[40] The consequences of the GDPR provisions about informed consent on the scientific research context are still a highly debated issue.[41]

Respecting the consent requirements listed by the GDPR could be problematic in the field of predictive medicine. Firstly, the functioning of AI algorithms is unintelligible for human observers: Even if there is a clear outcome from the working process of the system, the reasoning pattern remains obscure. Thus, the data subject cannot understand how their personal data are collected and processed: Consent could not be defined as 'informed' as required by the GDPR. Secondly, consent is not even 'specific' because AI algorithms usually follow adaptive patterns to perform their interpretative tasks, changing their working routine in light of new circumstances. It is therefore not possible for data subjects to know all the specific features of the processing activities when they provide consent. This also prevents people from giving their approval freely, considering all potential consequences, or meaningfully, taking into account all possible variables and suitable alternatives.[42]

The development of machine learning systems requires a rethinking of the legal category of consent. It is necessary to transcend the traditional paradigm of consent, focused on a single specific purpose, to find a new legal solution compatible with the inherent features of AI working routine. Two

---

39 Article 29 Working Party, Guidelines on Transparency under Regulation 2016/679 (11 April 2018) 17/EN WP260 rev.01 <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051> accessed 2 March 2020.

40 On the issue of capacity to consent, see Michelle Biros, 'Capacity, Vulnerability, and Informed Consent for Research' (2018) 46 The Journal of Law, Medicine & Ethics 72

41 Chassang (n 32); Niam Clarke et al., 'GDPR: An Impediment to Research?', (2019) 188 Irish Journal of Medical Science 1129; Miranda Mourby and others, 'Governance of Academic Research Data under the GDPR – Lessons from the UK' (2019) 9 International Data Privacy Law 192.

42 Article 29 Working Party, Guidelines on Consent under Regulation 2016/679 (10 April 2018) 17/EN WP259 rev.01 <https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=623051> accessed 2 March 2020.

approaches are worth briefly mentioning. Firstly, the broad consent[43] model, usually applied in the context of biobanks, informs data subjects about the overall scope and modalities of the data processing activities but not the specific processes behind these procedures. Secondly, the dynamic consent[44] solution establishes a constant dialogue, through a digital platform, between data subjects and controllers, allowing patients to understand how their data is processed in successive diagnostics operations and exercise continuous control over the processing of their personal data, including by withdrawing their previous consent.

IV. The Right to Be Forgotten: How Can an AI Algorithm Forget its "Memory"?

AI algorithms' need to train datasets to improve their processing capabilities could raise problems with one of the most relevant innovative features introduced by the GDPR: the so-called right to be forgotten.[45] Formulated by the Court of Justice of the European Union in the famous judgement *Google Spain,* it recognizes the data subject's right to obtain from the controller the erasure of personal data concerning them without undue delay.[46] The right to be forgotten applies to different circumstances enumerated by the GDPR itself, such as when data are no longer needed for the original purposes, or when data subjects withdraw their consent. Where an individual exercises their right to be forgotten, the data controller must take reasonable measures to erase the data from the public domain, also removing any links related to them.

From a practical perspective, the inherent technological features of AI algorithms may complicate the application of the right to be forgotten within the field of predictive medicine. Physicians feed medical data to the algorithms to train the computer programmes, which acquire new

---

43 Mark A. Rothenstein, 'Broad Consent for Future Research: International Perspectives' (2018) 40(6) Ethics & Human Research 7

44 Jane Kaye and others, 'Dynamic Consent: A Patient Interface For Twenty-First Century Research Networks' (2014) 23 European Journal of Human Genetics 141.

45 GDPR, art 17.

46 Case C-131/12, *Google Spain SL, Google Inc. v Agencia Española de Protección de Datos, Mario Costeja González* [2014] ECR 317.

information to increase the overall knowledge of algorithms. Thus, the data to be erased are no longer a separate unit, but part of the AI software experience. As a result, it would be technologically impossible to extract a single piece of data without interfering with the reasoning process of the algorithm. Removing data from the AI system would radically change in production of outcomes, potentially harming patients.

V. The Prohibition of Automated Decision-making and Profiling Activities under the GDPR as a Regulatory Challenge for the Development of Black-box Medicine

AI algorithms can produce outcomes autonomously, or at least with minimal involvement of human observers. This raises ethical and legal concerns about safeguarding the fundamental rights of people subject to the action of automated processing activities. Specifically, the GDPR grants the data subject the 'right 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'.[47] The main goal is to prevent the outcomes of decision-making algorithms from infringing people's fundamental rights,[48] as machine and computer systems can formulate decisions based on inaccurate or harmful data sets that yield a misleading or biased interpretation of reality.[49]

---

47 GDPR, art 22.1

48 Article.22.3 of the GDPR prescribes that the data controller shall provide 'suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision'. More specifically, Recital 71 of the GDPR explains that such safeguards should include 'specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision'. Gianclaudio Malgieri, 'Automated Decision-Making in the EU Member States: The Right to Explanation and "Suitable Safeguards" in the National Legislations' (2019) 35 Computer Law & Security Review 105327.

49 Milena A.Gianfrancesco and others, 'Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data' (2018) 178 The Journal of American Medical Association 1544.

Under the GDPR, automated decision-making activities based on sensitive data are unlawful without the prior explicit consent of the data subject except in cases of overriding public interest.[50] On its face, this prohibition purports to apply only to decisions taken without any human intervention whatsoever. Since the working routine of AI tools still often requires some sort of external action, such a restriction would apply only very rarely. Thus, the precise scope of this prohibition is open to interpretation; perhaps it applies only to decisions made by the algorithm without a *meaningful* human involvement.[51]

The general prohibition of decisions based solely on automated processes could deter the development of black-box medicine. Humans have only a secondary role in black-box medicine. Physicians make treatment decisions by considering the outcomes produced by the AI algorithms, but cannot replicate the reasoning process of the machine. Limiting the use of AI programmes to cases of previous explicit consent or overriding public interest is too narrow in scope. It is crucial to find an appropriate balance between the right to privacy and data protection of every individual and the use of innovative tools to guarantee higher health standards for the entire community. The challenge is to set boundaries between the right to health and the protection of personal data.

The GDPR could provide valuable indications to minimize doubt. It already recognizes that the protection of personal data is not an absolute right, but rather 'must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality'.[52] It already sanctions processing activities 'carried out in the public interest'.[53] It is unquestionable that the right to public health is an issue of public interest. Nonetheless, fundamental safeguards to privacy rights must be preserved even in the functioning of AI algorithms for healthcare.

---

50 GDPR, art 22.4.

51 Gianclaudio Malgieri and Gianni Comandè, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 243.

52 GDPR, recital 4.

53 GDPR, art 6.4.

Several proactive and preventive data protection measures in AI algorithms could adequately safeguard the privacy rights of patients involved in diagnostic treatments. Firstly, a counterfactual explanation model could allow individuals to better understand the reasoning process behind a predictive outcome.[54] This would explain what factors would need to change to obtain a different result, permitting scientists and physicians to understand the relationship between processed data and the above-mentioned principles of data processing. Secondly, a co-governance system of algorithms based on a multi-level design could allow humans to intervene in the reasoning patterns to ensure the respect of fundamental rights of the patients involved.[55] Thirdly, the so-called 'agonistic machine learning'[56] approach, where AI providers should formulate alternative ways of modelling and describing the same object, could provide new diagnostic patterns compliant with the fundamental rights framework. AI algorithms usually rely on machine-readable information about what is the 'truth': for instance, in the health care sector, medical exams or diagnoses. Providing different inputs from several sources would help algorithms overcome possible biases in the datasets and produce more reliable outcomes. This may lead to a more accountable and transparent decision-system, complying with the data protection framework.

VI. Privacy by Design and Privacy by Default in the Era of AI Algorithms

The GDPR requires compliance with the principle of data minimisation,[57] whereby controllers must process only the necessary amount of information. The principle of privacy by design, developed by Ann Cavoukian,[58] is a proactive approach to data minimisation, integrating privacy measures into

---

54 Sandra Wachter and others, 'Counterfactual Explanations Without Opening the Black Box' (2018) 31 Harvard Journal of Law & Technology 841.

55 Margot E. Kaminski, 'The Right to Explanation, Explained' (2019) 34 Berkeley Technology Law Journal 189.

56 Mireille Hildebrandt, 'Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning', (2019) 19 Theoretical Inequalities of Law 83.

57 GDPR, art 5.

58 Cavoukian (n 10).

the hardware and software utilised in data processing upon their creation to ensure that only the necessary amount of data is processed. Privacy by design requires data controllers to implement adequate privacy safeguards, such as pseudonymisation, from the first phases of the processing activities.[59]

This principle could conflict with the inherent nature of AI algorithms. AI tools constantly need data to train their working routine to perform their diagnostic tasks more efficiently.[60] Furthermore, algorithms adapt to constantly changing environments; maintaining the same data protection features may be technically impossible, though technical approaches like counterfactual explanations, co-governance systems of algorithms or agonistic machine learning could help bridge the gap.

Ultimately, the working of machine learning tools should be considered compliant with the GDPR provisions to ensure a normative safeguard for the rights of data subjects against the risk of obsolescence of the Regulation. The privacy by design principle could play a fundamental role to avoid this kind of risk and keep the pace of technological progress. This proactive and preventive approach would make the user—the patient in the medical setting—the focus of the entire data processing activity. Embedding privacy issues in the construction of AI algorithms would also help to keep track of the reasoning patterns chosen to produce a specific output. Privacy by design would encourage dialogue between AI providers, scientists and privacy advocates to build privacy-compliant AI algorithms that could help physicians and scientists manage the risks related to processing health data, taking into account the privacy rights of people involved.

VII. Concluding Remarks

In a time of new approaches to data protection, the GDPR remains the 'gold standard' in the European framework. However, the GDPR reflects an era when AI had not yet reached the current levels of technological development and, more specifically, the field of predictive medicine was in its infancy. As a result, the Regulation is not fully compatible with the inherent features of

---

59 GDPR, art 25.

60 European Parliamentary Research Service, *The Impact of the General Data Protection Regulation (GDPR) on Artificial Intelligence*, (European Union 2020).

machine learning tools. The resulting legal uncertainty could osbstruct the development of AI technologies, increasing costs and reducing benefits for users and patients.[61] The EU must act to bridge this gap and fully regulate the use of AI tools in everyday life, including the medical sector, and achieve a uniform and coherent policy approach regarding AI matters. Encouragingly, the European Commission has acknowledged the threats to privacy posed by machine learning applications and cleared the way for adjusting relevant EU legislative frameworks.[62]

GDPR norms are often vague and undefined,[63] which may be a normative choice to keep pace with the ongoing technological progress. Specific mandatory requirements aimed at AI tools may become rapidly obsolete. It is necessary instead to create a trustworthy environment for developing AI applications for healthcare purposes with patient privacy rights in mind. The GDPR already provides valuable instruments, such as the privacy by design principle,[64] that could help reach this goal. Embedding data protection features in diagnostics routines would help overcome the black-box barrier of algorithms. Software providers would train their AI tools to respect privacy rules from the very first phases of their working patterns, securing lifecycle protection for the user during the entire duration of data processing activities. The patient would become the focus of the diagnostic process. Scientists and physicians would coordinate the work of AI algorithms. Humans would control the entire process; not machines. This would ensure the full respect of human rights of every individual involved, resolving the tension between the privacy rights of the individual and the public health needs of society.

---

61 Ibid.

62 European Commission, 'White Paper on Artificial Intelligence – A European Approach to Excellence and Trust' COM (2020) 65 final.

63 Ibid.

64 Cavoukian (n 10).