European Journal of Risk Regulation (2024), **15**, 903–911
doi:10.1017/err.2024.74

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Cybersecurity Framework for Synthetic Data in Training Medical AI

Jarosław Greser

Faculty of Administration and Social Science, Warsaw University of Technology, Cyber & Data Security Lab, Vrije Universiteit Brussels
Email: jaroslaw.greser@pw.edu.pl

Abstract

The development of medical artificial intelligence is dependent on the availability of vast quantities of data, a considerable proportion of which is medical data containing sensitive information pertaining to the health and well-being of patients. The use of such data is subject to extensive legal regulation and is further hindered by financial and organisational constraints, which can result in limitations on accessibility. One potential solution to this problem is the use of synthetic data. This article examines the potential for their use in light of cybersecurity requirements derived from horizontal and sectoral EU legislation. The outcome of this analysis is that EU legislation does not contain specific regulations on the use of synthetic data. Consequently, it cannot be concluded that there is any prohibition on their use. Moreover, while the Medical Device Regulation (MDR) contains some general requirements for cybersecurity, these are further specified by the provisions of the AI Act. It is important to note, however, that the AI Act will not apply to Class I medical devices, which are subject only to the MDR. Furthermore, only indirect obligations within the scope under consideration can be derived from the horizontal regulations, which will apply in a limited number of cases.

Keywords: cybersecurity; medical AI; medical devices; synthetic data

I. Introduction

The development of AI-based systems requires access to both large amounts of and high-quality data. This is particularly important for systems based on supervised and unsupervised machine learning. It is vital to note that these solutions account for the majority of deployments in the medical AI market, which is estimated to be worth $22.45 billion by 2023[1] and is projected to grow to $164.10 billion by 2029, at a compound annual growth rate of 42.4%.[2] However, medical data is subject to various restrictions

This article is the result of the research carried out in Laboratory for Data Protection & Privacy Impact Assessments (d.pia.lab) during research stay at Vrije Universiteit Brussels from January to June 2023.

[1] Grand View Research, "AI In Healthcare Market Size, Share & Trends Analysis Report By Component (Software Solutions, Hardware, Services), By Application (Virtual Assistants, Connected Machines), By Region, And Segment Forecasts, 2024–2030" (2023) <https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-healthcare-market> accessed 1 December 2023.

[2] Fortune Business Insights, "Artificial Intelligence in healthcare" (2022) <https://www.fortunebusinessinsights.com/industry-reports/artificial-intelligence-in-healthcare-market-100534> accessed 1 December 2023.

when used to train medical AI systems including legal,[3] ethical[4] and organisational factors.[5] This can be illustrated by the example of avoiding bias. In practice, this problem arises, for example, when databases are created from merging patients' data and public data, which is common practice.[6] This helps to create models with high relevance – especially for image generation or natural language processing.[7] However, it is difficult to control their content,[8] so if biased data is used for training, model will be biased.

One of the solutions to the above-mentioned issues may be the use of synthetic data.[9] Currently, there is a discussion on the possibilities and conditions for its application in different sectors[10] including the medical one.[11] This article contributes to this discussion by identifying the legal requirements of cybersecurity as one of the bases for risk assessment when using this data to train medical AI systems. It consists of four parts. The first discusses what synthetic data is and its prospects for use in the medical sector. The second focuses on the cybersecurity vulnerabilities of AI systems. The third presents the legal requirements for training medical AI systems from a cybersecurity perspective. The fourth concludes with assessing the feasibility of using synthetic data to train medical AI systems and making recommendations in this regard.

II. Synthetic data and its use in the medical sector

There is no legal definition of synthetic data. However, this concept is widely recognised in the technical literature, which makes it possible to establish factors distinguishing it from other types of data. Generally, two aspects have been highlighted in the studies. The first is its source: such data sets are created rather than collected. This has various consequences, among which is the conclusion that these sets will always be artificial in the sense that there are no equivalents in the "real" world. Moreover, a generating algorithm is required to create such data. It is underlined that many different methods

[3] Elisabetta Biasin, Burcu Yaşar, Erik Kamenjašević, "New Cybersecurity Requirements for Medical Devices in the EU: The Forthcoming European Health Data Space, Data Act, and Artificial Intelligence Act" (2023) 5(2) Law, Technology and Humans 43.

[4] Sarah E Hickman, Gabrielle C Baxter, Fiona J Gilbert, "Adoption of artificial intelligence in breast imaging: evaluation, ethical constraints and limitations" (2021) 125 British Journal of Cancer 15.

[5] Cristina Trocin, Patrick Mikalef, Zacharoula Papamitsiou, Kieran Conboy, "Responsible AI for Digital Health: a Synthesis and a Research Agenda" (2023) 25 Information Systems Frontiers 2139.

[6] Natalia Ponomareva, Jasmijn Bastings, Sergei Vassilvitskii, "Association for Computational Linguistics Training Text-to-Text Transformers with Privacy Guarantees" (2022) <https://aclanthology.org/2022.findings-acl.171.pdf> accessed 1 December 2023.

[7] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, et al., "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer" (2020) 21(140) Journal of Machine Learning Research <https://jmlr.org/papers/v21/20-074.html> accessed 1 December 2023.

[8] Anna Rogers, "Changing the World by Changing the Data" (2021) Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing <https://aclanthology.org/2021.acl-long.170.pdf.> accessed 1 December 2023.

[9] Erroll Wood, Tadas Baltrušaitis, Charlie Hewitt, Sebastian Dziadzio, Matthew Johnson, Virginia Estellers, et al., "Fake It Till You Make It: Face analysis in the wild using synthetic data alone" (2021) <https://arxiv.org/abs/2109.15102> accessed 1 December 2023.

[10] Jiri Hradec, Massimo Craglia, Margherita Di Leo, Sarah De Nigris, Nicole Ostlaender, Nicholas Nicholson, *Multipurpose synthetic population for policy applications* (Publications Office of the European Union, Luxembourg 2022) 1, 15–16.

[11] Trivellore E. Raghunathan, "Synthetic data" (2021) 8 Annual Review of Statistics and Its Application 129

may be employed to create synthetic data,[12] and that they may be based on different learning methods[13] with their own advantages and disadvantages.[14]

The second distinction is the relationship of the data to the real world. Synthetic data is a statistical reflection of the properties of the original set, which in most cases will be real-world data. In theory, data scientists should draw the same statistical conclusions from analysing a given set of synthetic data as they would from real data. In practice, synthetic data is statistically relatable to the set from which it was created. This does not mean that they accurately reflect the reality. How accurately they reflect the reality depends on the database from which it is generated, and such a selection depends on the creator of the synthetic data. Naturally, the objection can also be raised against real-world databases that they do not reflect the "real world" in epistemological terms. However, in the case of synthetic data, explaining the accuracy may be much more complicated and lead to errors.

There are many positive aspects to the generation of synthetic data, especially when this data contains medical information. The first is a significant reduction in the cost of preparing the database. This includes cleaning, labelling and organising the raw data sets. For example, data can be extracted from the electronic medical record (EMR) used in the hospital. It contains different types of data, in particular consultations and diagnostic data, but it can be much broader and include pharmacy prescriptions, insurance records, genomics-driven experiments such as genotyping or gene expression data.[15] It may also include automatically collected data from the Internet of Things (IoT).[16] In addition, healthcare professionals belong to different sectors, such as dentistry, medicine, nursing or physiotherapy, which may result in data input according to different methodologies. There is also a problem with data interoperability, especially when it comes from different medical facilities,[17] and the common practice has been to keep part of the documentation in the form of either handwritten notes or typed reports.[18] All these factors result in the situation where transforming EMRs into high-quality databases for AI training purposes can be very resource consuming. In the case of synthetic data, generating it according to a specific algorithm makes it structured according to a specific key.

Another issue is the ability to use synthetic data to easily increase the database variety in cases where access to patients is limited. This is a major challenge when the patient population is limited in number, as in the case of rare diseases, or when the ability to test is limited due to lack of patient consent, as in the case of pregnant people or children, or due to recruitment problems, which often occur in the case of disadvantaged groups.[19] The literature suggests that underrepresented datasets may be biased[20], and their use may lead to erroneous results and violations of fundamental rights[21]. One of the solutions to this

[12] Hradec (n 8) 12–15.

[13] Atijit Anuchitanukul, Julia Ive, "SURF: Semantic-level Unsupervised Reward Function for Machine Translation" (2022) <https://aclanthology.org/2022.naacl-main.334.pdf> accessed 1 December 2023

[14] Hradec (n 8) 8.

[15] Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma, Sandeep Kaushik, "Big data in healthcare: management, analysis and future prospects" (2019)6 Journal of Big Data 54.

[16] Jarosław Greser, "Etyczne problemy wdrażania medycznego Internetu Rzeczy" (2020) 3 Prawo Mediów Elektronicznych 4.

[17] Miriam Reisman, "EHRs: the challenge of making electronic data usable and interoperable" (2017) 42(9) Pharmacology & Therapeutics Journal 572.

[18] Susan Doyle-Lindrud, "The evolution of the electronic health record" (2015) 19(2) Clinical Journal of Oncological Nursing 153.

[19] Rebeca Dresser, "Wanted. Single, White Male for Medical Research" (1992) 22 Hastings Center Report 1.

[20] Carsten Schwemmer, Carly Knight, Emily Bello-Pardo, Stan Oklobdzija, Martijn Schoonvelde, Jeffrey Lockhart, "Diagnosing Gender Bias in Image Recognition Systems" (2020) 6 Socius 1.

[21] Joy Buolamwini, Timnit Gebru, "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification" (2018) 81 Proceedings of Machine Learning Research 1.

problem is the use of synthetic data, especially in areas where it is highly relevant, such as image generation or natural language processing.[22]

A frequently raised argument for the use of synthetic data in medicine is that it ensures patient privacy. This is an essential element of its use from a legal and ethical perspective, but it can pose a major challenge in practice.[23] The proponents point out that synthetic data must be considered as anonymised data and as such is not subject to data protection regulations.[24] Research shows that such a claim can be true, but only in specific cases and when additional conditions are met.[25] It will also not be possible to treat it as pseudo-anonymised data in all cases, in particular if it shows sufficient structural equivalence to the original dataset or share relevant properties or patterns that could lead to the attribution of information to an individual.[26] This leads to the conclusion that the current state of the art does not allow synthetic data to be regarded as synonymous with anonymised data. Therefore, it can be assumed that the data protection legislation does apply to them. Even if we consider such data to be pseudo-anonymised, it is clear from the GDPR that technical and organisational measures must be taken to protect such data. Therefore, it can be assumed that in the case of synthetic data for medical AI training, a very careful analysis is required of the level of privacy offered by the collection and the risk of violating the rights of the individuals whose data were used to create the dataset. This does not preclude the use of synthetic data to train medical AI systems, but it does limit its use due to the boundaries imposed by data protection requirements.

III. Cybersecurity vulnerabilities of AI systems

Like any IT system, AI is vulnerable to cyber threats. In the case of AI, the European Union Cyber Security Agency lists dozens of threats classified in eight main areas.[27] These can be divided into two main groups. The first are threats that affect all ICT systems, such as the theft of information, preventing authorised users from accessing data, or unauthorised modifications of data in the system. In this case, countermeasures are relatively well known and described.[28]

The second group are vulnerabilities specific to artificial intelligence. The most serious are data poisoning and adversarial attacks. The former is a type of attack where data or a

[22] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, et al., "Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer" (2020) 21(140) Journal of Machine Learning Research <https://jmlr.org/papers/v21/20-074.html> accessed 1 December 2023.

[23] Hao Jin; Yan Luo; Peilong Li; Jomol Mathew, "A review of secure and privacy-preserving medical data sharing" (2019) 7 IEEE Access <https://ieeexplore.ieee.org/abstract/document/8713993/> accessed 1 December 2023

[24] Stylianos Kampakis, "How to beat GDPR in research: synthetic data" <https://thedatascientist.com/gdpr-research-artificial-data> accessed 1 December 2023.

[25] Julia Ive, "Leveraging the potential of synthetic text for AI in mental healthcare," (2022) 4 Frontiers Digital Health, Sec. Digital Mental Health; Theresa Stadler, Bristena Oprisanu, Carmela Troncoso, "Synthetic Data – Anonymisation Groundhog Day" (31st USENIX Security Symposium, Boston August 2022) <https://www.usenix.org/conference/usenixsecurity22/presentation/stadler> accessed 1 December 2023.

[26] César Augusto Fontanillo López and Abdullah Elbi, "On synthetic data: a brief introduction for data protection law dummies" (*European Law Blog* 22 September 2022) <https://europeanlawblog.eu/2022/09/22/on-synthetic-data-a-brief-introduction-for-data-protection-law-dummies> accessed 1 December 2023.

[27] ENISA, *Artificial Intelligence Cybersecurity Challenges. Threat Landscape for Artificial Intelligence*, (2020) <https://www.enisa.europa.eu/publications/artificial-intelligence-cybersecurity-challenges/@@download/fullReport> accessed 1 December 2023.

[28] Lynne Coventry, Dawn Branley, "Cybersecurity in healthcare: A narrative review of trends, threats and ways forward" (2018) 113 Maturitas 48; Khanyisile Vilakazi, Funmi Adebesin, 'Cybersecurity Threats to Healthcare Data and Mitigation Strategies' (2023) 93 EPiC Series in Computing 240.

model is altered to change the behaviour of an algorithm in a way that the attacker intends.[29] For example, instructing the algorithm that images of cancer represent a healthy tissue so that a similar image will be interpreted similarly in the future. Such attacks can occur at most stages of the project lifecycle, but the data collection and training stages of the algorithm are particularly vulnerable.[30]

An adversarial attack consists of a small change to the algorithm's input data that causes machine learning models to misclassify examples that are only slightly different from the correctly classified examples. Consequently, there are significant changes in the results obtained, leading to decision errors[31]. For example, a change of one pixel in the image of a frog leads to the image being misclassified as a dog or a truck[32]. The effects of this attack usually occur during the last lifecycle of a project, i.e. during its practical application, and therefore are relatively difficult to detect. The task is also made more difficult by the fact that, in the case of images, it is essentially impossible to point out images that have been deliberately altered by a human.[33]

Both types of attack have been reported by AI researchers for many years.[34] The literature describes their various taxonomies, methods of use and countermeasures.[35] System using synthetic data seems particularly vulnerable as these attacks can take place at any stage of a system's lifecycle.[36] Moreover, data poisoning and adversarial attacks are characterised by their high level of effectiveness.[37] From a medical AI perspective, they are considered particularly dangerous because a successful attack can result in a life-threatening or fatal outcome for the patient.[38] This risk is exacerbated by the difficulty in finding effective ways to defend against this type of attack, in part due to the failure to reduce the risk of attack when training the algorithm on inconsistent training data, and the lack of correlation between the explainability of the algorithm and the effectiveness of the attack.[39]

IV. The legal requirements for cybersecure training of medical AI systems

As stated above, there is no legal regulation that specifically addresses synthetic data, and therefore legal requirements will need to be reconstructed from regulations governing the cybersecurity of medical AI. The crucial issue in this regard is the distinction between AI that is

[29] ENISA, *Securing Machine Learning Algorithms*, (2021) 14 <https://www.enisa.europa.eu/publications/securing-machine-learning-algorithms/@@download/fullReport> accessed 1 December 2023.

[30] Ankush Mitra, Basudeb Bera, Ashok Kumar Das, Sajjad Shaukat Jamal, Ilsun You, "Impact on blockchain-based AI/ML-enabled big data analytics for Cognitive Internet of Things environment" (2023) 197 Computer Communications 173 <https://doi.org/10.1016/j.comcom.2022.10.010.> accessed 1 December 2023

[31] ENISA (n 27) 13.

[32] Naveed Akhtar, Ajmal Mian, "Threat of Adversarial Attacks on Deep Learning in Computer Vision: A Survey" (2018) 6 IEEE Access 14410.

[33] However, it is pointed out that it is possible to mislead the algorithm with images that human easily distinguish as manipulation. Ibid, 1–2.

[34] Christian Szegedy, Wojciech Zaremba, Ilya Sutskever, Joan Bruna, Dumitru Erhan, Ian Goodfellow, Rob Fergus, "Intriguing properties of neural networks" (2012) arXiv:1312.6199 <https://arxiv.org/abs/1312.6199> accessed 1 December 2023.

[35] Piotr Biczyk, Łukasz Wawrowski, "Towards Automated Detection of Adversarial Attacks on Tabular Data," (18th Conference on Computer Science and Intelligence Systems (FedCSIS), Warsaw, September 2023).

[36] James Jordon, Lukasz Szpruch, Florimond Houssiau, Mirko Bottarelli, Giovanni Cherubin, Carsten Maple, et. al., "Synthetic Data – what, why and how?" (2022) The Royal Society <https://arxiv.org/pdf/2205.03257.pdf> accessed 1 December 2023.

[37] Jiawei Su, Danilo Vasconcellos Vargas, Sakurai Kouichi, One pixel attack for fooling deep neural networks, (2017) arXiv:1710.08864 <https://arxiv.org/abs/1710.08864> accessed 1 December 2023

[38] Marco Eichelberg, Klaus Kleber, Marc Kämmerer, "Cybersecurity Challenges for PACS and Medical Imaging" (2020) 27(8) Academic Radiology 1126.

[39] Goodfellow (n 32) 1.

a medical device and AI that does not fall into this category. In the latter case, although the impact of such solutions on the market may be significant, they cannot in principle be used by healthcare professionals.[40] For this reason, I will exclude them from further consideration.

In the case of medical devices, we can look for solutions in sectoral and horizontal cybersecurity regulations and data protection legislation. Among the sectoral regulations, the Medical Device Regulation (MDR) and the In Vitro Device Regulation (IVDR) will play an important role. These regulations are based on the idea that a device can only be placed on the market or put into service if it complies with the general safety and performance requirements set out in the regulations. This includes, in particular, compliance with the general safety and performance requirements set out in Annex I. In addition, depending on the class to which the device is assigned, the legislation may impose additional requirements, such as the implementation and maintenance of a risk management system, the conduct of a clinical evaluation of the device, including post-market surveillance, or the preparation and updating of technical documentation.

AI solutions can be classified as medical devices, both as a stand-alone algorithm and as part necessary for the functioning of the device.[41] In the EU, the MDR/IVDR do not specifically mention cybersecurity or AI based on the assumptions that underpin the approach to the medical device regulation in the EU.[42] This means that AI solutions are subject to the same rules as other medical devices. However, the regulations contain provisions for electronic programmable systems, understood as devices containing electronic programmable systems and software, which are devices in their own right. According to paragraph 17 of Annex 1, such solutions must be designed and manufactured in accordance with the state of the art taking into account the principles of development life cycle, risk management, including information security, verification and validation and ensure repeatability, reliability and performance in line with their intended use. In a similar vein, general guidelines are provided by the Medical Devices Coordination Group. They highlight the need for security by design, security verification and validation testing, and security update management, but do not address the specific requirements or risks of artificial intelligence technologies.

A sector-specific legislation that may be relevant to the regulation of cybersecurity requirements for medical AI is the AI Act. The proposal[43] included a solution to consider medical devices as high-risk systems, but such a qualification would not mean that the algorithm would be considered high-risk under the MDR/IVDR. At the same time, under Article 47, medical devices would not be subject to an additional conformity assessment procedure and notification of serious incidents or malfunctions will be limited to those that constitute a breach of obligations under European Union law intended to protect fundamental rights. The Council's General Approach[44] reinforces the concept of imposing the requirements of the AI Act on medical AI. Recital 54a states that the AI Act "should apply without prejudice to more specific provisions laid down in certain sectoral legislation of the New Legislative Framework with which this Regulation should apply jointly." In addition, Article 6(1) states that an AI system which is itself a product covered by European Union harmonisation legislation shall be considered as high risk if it is subject

[40] B. Mittelstadt, "Ethics of the health-related internet of things: a narrative review," (2017) 19(3), Ethics and Information Technology, 157.

[41] Anastasiya Kiseleva, "AI as a Medical Device: Is It Enough to Ensure Performance Transparency and Accountability in Healthcare?" (2020) 1 European Pharmaceutical Law Review.

[42] Elisabetta Biasin, Erik Kamenjašević, "Regulatory Approaches Towards AI-Based Medical Device Cybersecurity: an EU/US Transatlantic Perspective" (2024) 1 European Journal of Risk Regulation (forthcoming)'.

[43] Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts [2021] COM(2021) 206 final 2021/0106(COD).

[44] Council's General Approach, 6 December 2022 2021/0106(COD).

to third party conformity assessment for the placing on the market or putting into service of that product in accordance with that legislation. This does not mean that an AI system will not automatically be considered "high risk" under the MDR/IVDR, but that for medical systems it will need to meet the additional requirements that the AI Act provides.

As regards horizontal legislation, the most important are the NIS and NIS2 Directives, which aim to create a legal framework for the development of national cybersecurity systems and networks for information exchange and cooperation between EU countries. The former, which is currently in force, imposes obligations on essential and important entities. Therefore, the mere fact of the implementation of any type of medical AI does not automatically bring it within the scope of the Directive. It is possible, however, that a decision by a Member State may confer on such an entity a status that becomes the source of its obligations.[45] Such an arrangement has led to differences in interpretation and thus in implementation in the Member States.[46] This led to regulatory work culminating in the adoption of the NIS2 Directive, which is due to be implemented in Member States by October 2024. Under this provision, entities manufacturing medical devices as defined in Article 2(1) of Regulation (EU) 2017/745 are qualified as "important entities," and consequently all the associated obligations apply to them. Furthermore, entities manufacturing medical devices that are considered critical during a public health emergency as defined in Article 22 of Regulation 2022/123 will be considered as "essential entities." It should be noted that these entities will be subject to additional obligations under the CER Directive,[47] which explicitly mentions its application in point 5 of the Annex.

Another horizontal piece of legislation is the Cybersecurity Act. The European Cybersecurity Certification Framework, based on Article 46 of this Act, may refer directly to medical devices. In the case of medical AI, there are currently no such schemes, but it is possible that they will emerge in the future. For the time being, however, this possibility remains theoretical.

The final group of regulations that may impose requirements on the training of medical artificial intelligence are data regulations. These fall into two groups. The first is legislation aimed at implementing the European Data Strategy adopted in 2020.[48] These include legislation on the creation of data spaces[49] or the harmonisation of data access rules.[50] Studies on their impact on cybersecurity are available in the literature,[51] but as they are at the stage of legislative initiatives, I will not discuss them further due to possible major changes in the final version.

With regard to data protection legislation, the most important role is played by the GDPR, in particular Articles 5(1)(f) and 32. Given that personal data is a concept that is

[45] Medical Device Coordination Group, "MDCG 2019-16 Guidance on Cybersecurity for Medical Devices" (2019) 33 <https://health.ec.europa.eu/system/files/2022-01/md_cybersecurity_en.pdf> accessed 1 December 2023.

[46] Communication from the Commission of 10.3.2020 to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on the definition and removal of barriers in the single market, COM(2020) 93 final.

[47] Council Directive 2022/2557 of 14 December 2022 on the resilience of critical entities and repealing Council Directive 2008/114/EC, OJ L333/164.

[48] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions of 19.2.2020 'A European Data Strategy', COM(2020) 66 final.

[49] Proposal of 23.2.2022 Regulation of the European Parliament and of the Council on the European Health Data Space, [2022] COM/2022/197 final.

[50] Proposal of 23.2.2022 Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), [2022] COM(2022) 68 final.

[51] Biasin, Elisabetta, Burcu Yaşar, and Erik Kamenjaševic, "New Cybersecurity Requirements for Medical Devices in the EU: The Forthcoming European Health Data Space, Data Act, and Artificial Intelligence Act" (2023) 5 (2) Law, Technology and Humans 43.

interpreted very broadly,[52] and that when personal and non-personal datasets are linked, the GDPR rules must be applied to all data,[53] it is reasonable to assume that the requirements of this legislation will apply to the vast majority of datasets used to train medical AI.

V. Conclusions

Synthetic data has many advantages that make it potentially useful for training artificial intelligence systems. At the same time, using it for this purpose introduces an additional layer that needs to be considered when analysing cybersecurity risks. It should be noted that the threat of an attack on medical AI is real. Although no such incident has been reported to date, given that medical infrastructure is one of the main targets of cyber attacks, it is reasonable to assume that such an attack will occur at some point. One way to mitigate the risk is to comply with legal requirements. Looking at legislation at the EU level, there is a mosaic of regulatory requirements that have their sources in different pieces of legislation. It includes regulations governing putting on the market of medical devices, norms on artificial intelligence, and horizontal regulations on cybersecurity. Several conclusions can be drawn from this.

The first conclusion is that there is no general prohibition on the use of synthetic data to train medical AI systems. However, it should be pointed out that this results more from the fact that the issue is relatively new and its legal implications are only now being analysed. Nevertheless, according to the principle "quod lege non prohibitum, licitum est" (what is not forbidden by law is allowed), producers of medical AI can use this data. At the same time, they will have to take full responsibility for cybersecurity of system they produce and put on the market.

A further argument in favour of the possibility of using synthetic data in training medical AI algorithms is the assumption of technological neutrality of legal acts concerning the regulation of digital technologies, including cybersecurity.[54] According to this assumption,[55] the deployment of technological solutions unknown at the time of the adoption of the legislation is covered by the obligations that follow from it. In practice, it involves applying to legal text a risk-based approach to these solutions and formulating general obligations to design and manufacture them according to the state of the art. It should be emphasised that the legal acts analysed in the framework contain such clauses. Consequently, they can be used as a basis for setting the boundaries for cybersecure training of medical AI algorithms based on synthetic data.

To conclude, using synthetic data to train medical AI requires a very detailed recognition and assessment of the risks involved. Meeting the legal requirements that are imposed by the regulations governing the cybersecurity of medical devices gives an indication of the risks that need to be taken into account. However, it appears that meeting legal requirements may not be sufficient to effectively prevent attacks. Thus, manufacturers of medical AI should also take into account areas that for various reasons

[52] Nadezhda Purtova, "The law of everything. Broad concept of personal data and future of EU data protection law" (2018) 10 Law, Innovation and Technology 41.

[53] Christopher Kuner, Lee Brygave, Christopher Docksey, Laura Drechsler (eds), *The EU General Data Protection Regulation (GDPR). A commentary* (Oxford University Press 2020) 113.

[54] Teresa Rodríguez De Las Heras Ballell, "Legal Challenges of Artificial Intelligence: Modelling the Disruptive Features of Emerging Technologies and Assessing Their Possible Legal Impact," Uniform Law Review 24, no. 2 (June 1, 2019): 302–14.

[55] This assumption has been the subject of criticism from some authors, who have questioned its correctness in certain situations or highlighted its shortcomings. See: Bertolini, Andrea. "AI Does Not Exist! Defying the Technology-Neutrality Narrative in the Regulation of Civil Liability for Advanced Technologies." Europa e diritto privato (2022): n. pag. Print, p. 416.

are not regulated by law but are driven by industry standards. These include, for example, good practices in data collection and management or the creation of project documentation.[56] A comprehensive approach reduces risks of a different nature, which will positively influence the level of cybersecurity of the AI solution that synthetic data has been used to train.

[56] In the case of AI systems, there are voluntary initiatives that create documentation standards for algorithms proposed by both researchers, and public entities and some of these have been adopted widely by the AI community. Isabelle Hupont, Marina Micheli, Blagoj Delipetrev, Emilia Gómez, Josep Soler Garrido, "Documenting high-risk AI: A European regulatory perspective" (2022) TechRxiv. <https://www.techrxiv.org/articles/preprint/Documenting_high-risk_AI_an_European_regulatory_perspective/20291046> accessed 1 December 2023.

Cite this article: J Greser (2024). "Cybersecurity Framework for Synthetic Data in Training Medical AI". *European Journal of Risk Regulation* **15**, 903–911. https://doi.org/10.1017/err.2024.74