



The European Artificial Intelligence Strategy: Implications and Challenges for Digital Health

Cohen, Glenn ; Evgeniou, Theodoros; Gerke, Sara; Minssen, Timo

Published in:
The Lancet Digital Health

DOI:
[10.1016/S2589-7500\(20\)30112-6](https://doi.org/10.1016/S2589-7500(20)30112-6)

Publication date:
2020

Document version
Publisher's PDF, also known as Version of record

Document license:
[CC BY](#)

Citation for published version (APA):
Cohen, G., Evgeniou, T., Gerke, S., & Minssen, T. (2020). The European Artificial Intelligence Strategy: Implications and Challenges for Digital Health. *The Lancet Digital Health*, 2(7), 376-379.
[https://doi.org/10.1016/S2589-7500\(20\)30112-6](https://doi.org/10.1016/S2589-7500(20)30112-6)



The European artificial intelligence strategy: implications and challenges for digital health

I Glenn Cohen, Theodoros Evgeniou, Sara Gerke, Timo Minssen

In February, 2020, the European Commission published a white paper on artificial intelligence (AI) as well as an accompanying communication and report. The paper sets out policy options to facilitate a secure and trustworthy development of AI and considers health to be one of its most important areas of application. We illustrate that the European Commission's approach, as applied to medical AI, presents some challenges that can be detrimental if not addressed. In particular, we discuss the issues of European values and European data, the update problem of AI systems, and the challenges of new trade-offs such as privacy, cybersecurity, accuracy, and intellectual property rights. We also outline what we view as the most important next steps in the Commission's iterative process. Although the European Commission has done good work in setting out a European approach for AI, we conclude that this approach will be more difficult to implement in health care. It will require careful balancing of core values, detailed consideration of nuances of health and AI technologies, and a keen eye on the political winds and global competition.

Introduction

On Feb 19, 2020, the European Commission published a much-awaited white paper¹ on artificial intelligence (AI), and an accompanying communication² and report,³ fulfilling a major pledge⁴ in setting out a European approach for AI. The paper emphasises the need for the EU to “act as one and define its own way, based on European values, to promote the development and deployment of AI.”¹ For regulation and investment, the white paper presses a dual goal: support the AI uptake and address risks linked to particular uses of it.¹ The paper regards health as an important application area, and sets the lofty goal for the EU to “become a global leader in innovation in the data economy and its applications.”¹

The Commission's progress is commendable. Its views stay true to its European values as part of the strategy to address the increasing competition from the US, Chinese, and soon UK industries. It also identifies some of the unique challenges of AI for regulators and contemplates a regulatory structure that also evolves to fit these characteristics. The Commission recognises the value-laden nature of AI regulation and that Europeans have distinct views about privacy, transparency, fairness, etc. Nevertheless, as applied to medical AI, we argue that there are important challenges in the European Commission's approach that might be inimical if not fulfilled. We also outline what we view as the most relevant next steps in the Commission's iterative process.

European values and European data

The Commission emphasises “AI based on European values” and even calls on the EU to “export its values across the world.”¹ This lofty statement of purpose, runs into some immediate questions—if not obstacles. First, several important shared values are embedded in the legal structure of the Union itself and in key existing forms of EU legislation and initiatives. These values include, for example, the European Green Deal, the Charter of Fundamental Rights of the EU, and certain privacy

protections captured by the EU General Data Protection Regulation (GDPR). However, the Commission might be too optimistic that the questions raised by medical AI regulation will be resolvable based on well-established shared European norms and values. This approach, of course, reflects only part of the EU vision: that even with European values, there persists “the idea of a national constitutional identity that is reflected in national constitutional values.”⁵

In the context of health care, especially some existing rights, like the fundamental right of non-discrimination, for instance based on race, might not provide clear answers to pressing questions posed by medical AI. For example, is it enough to satisfy the fundamental right of non-discrimination to build AI systems to ignore variables such as race, or is that right violated if the system nonetheless produces disparate impact on various racial groups? What happens if including race as a variable actually produces more racially equitable outcomes?^{6,7} To use another example, how important is opacity in medical AI as opposed to accuracy— that is, should we be willing to accept more black box medical AI paired with rigorous demonstration of its care-improving effects, or should we demand a right to an explanation in all cases even at the expense of innovations that can be shown to improve care?⁸ We are not sure one would find high levels of homogeneity within the EU on these and a myriad of other specific questions raised by the regulation of medical AI. Although not only European but also universal values are desirable, especially in the light of the current pandemic, this goal will be even more challenging to achieve.

To put the point another way: many complex regulatory design questions, and more difficult values questions raised by medical AI exist (eg, potential trade-offs between investment and transparency, explainability and accuracy, etc.). The conflicts in values will make agreeing on a new European regulatory framework on trustworthy AI more difficult. Thus, such a process will likely be long and arduous, as we have seen with the GDPR. However,

Lancet Digital Health 2020;
2: e376–79

Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School, Harvard University, Cambridge, MA, USA (Prof I G Cohen JD, S Gerke Dipl-Jur Univ); INSEAD, Fontainebleau, France (Prof T Evgeniou PhD); and Centre for Advanced Studies in Biomedical Innovation Law, Faculty of Law, University of Copenhagen, Copenhagen, Denmark (Prof T Minssen Jur Dr)

Correspondence to:
Prof I Glenn Cohen, Harvard Law School, Harvard University, Cambridge, MA 02138, USA
igcohen@law.harvard.edu

we understand that such a framework is paramount to make sure that at least a broad consensus exists between the member states on high-risk AI applications and avoid the risk of a fragmented single market. However, with the speed of AI development and the competitive pressures from the US, China, and soon the UK, any GDPR-like process is likely too slow-going, whereas a rushed job might leave gaps that prove just as counterproductive.⁹

However, even if these challenges could be overcome, how exactly does one ensure that values are built-in across cultures and contexts? In medicine, data are the recorded experiences of patients and physicians, but as such, they carry with them the preferences, practices, policies (and biases) of those health-care encounters. There is no reason to believe AI trained on EU health data will produce good results abroad. To use the most obvious example, algorithms trained with data in elite settings, such as leading hospitals, might not recommend appropriate actions, such as treatments in lower-resource settings.¹⁰ Does retraining the AI for low-resource settings compromise European values (such as robust health-care rights) in the eyes of the Commission or is it instead context-sensitive? The opposite problem occurs when AI trained on non-EU data is used in EU settings. A striking example of how such use can lead to values being mismatched is the painful experience of trying to import US electronic health records to Denmark.¹¹ The Epic system was designed for US physician's offices and hospital systems mainly to enable them to bill more efficiently.¹¹ Thus, the system could not be untied from the US medical culture and easily translated to Danish health-care systems that are based on socialised medicine.¹¹ In the context of AI the results are hard to predict. In some cases, requiring retraining of non-European AI using European data might deprive European citizens of possibly better performing AI systems, for example, if these were trained using larger or more diverse datasets—is that an acceptable quality loss in the name of EU values?

The update problem

For some products—eg, drugs and medical devices—regulators typically review a product before it goes on the market. How can they determine when the updated AI behaves differently enough that a new review is needed? This can be called the “update problem.”¹² The Commission acknowledges the problem—identifying a number of risks, including cybersecurity ones, that might arise through self-learning during usage—but does not wrestle with what to do. One approach would be to lock the algorithm—so that it does not evolve over time and does not use new data to alter its performance¹³—at the moment of review, but this foregoes much of the value of such systems.¹² Consider an algorithm that analyses mammograms and makes recommendations as to breast cancer: if its training data is primarily based on

Caucasian women, allowing the algorithm to update as more data from European women of African descent are included would be desirable since breast density differs by race.¹² As Walter Gretzky, the father of the former Canadian ice hockey player Wayne Gretzky, used to say: “Go to where the puck is going, not where it has been.”¹² Thus, if you are aiming to hit a moving target (say, a rapid testing device for an evolving virus or antimicrobial resistance), it makes no sense to aim at where it was before, you want to aim where it is going. Instead of trying to lock down algorithms, regulators should focus on developing processes to continuously monitor, identify, and manage risks associated with these algorithms.¹²

This approach will involve setting up new systems and processes. For example, continuous stress testing could be carried out using, for example, simulations and robustness analysis.¹¹ Sentinel, the US Food and Drug Administration's (FDA's) monitoring system for medical products, could also be used to non-stop monitor AI systems.^{12,14} The FDA also aims to expand Sentinel to three coordinating centres (ie, Sentinel Operations Center, Community Building and Outreach Center, and Innovation Center) and thus will probably have the capacity to implement such AI monitoring.^{12,14} However, the performance of AI algorithms is sensitive not only to any changes of the data used as inputs, whenever the technology is used in different contexts but also to many human or organisational factors such as differences in skill levels or cultures across hospitals and regions.¹³ Moreover, implementing such a system in the specific European judicial and political context will not be easy. In a meeting on Common Data Models (CDM) and Real World Data (RWD) uses held at the European Medicines Agency (EMA) in December, 2017, the EMA acknowledged that the Sentinel system provides the FDA with a very high level of control, and noted that although it “cannot answer all questions...any of the areas of weakness relate more to the characteristics of the administrative claims databases themselves rather than to the system”.¹⁵ However, the report also pointed out that the Sentinel system is expensive and that although “Europe is fortunate with its national health-care systems which provide longitudinal ‘cradle to grave’ care and in some members states have provided a wealth of data for research...there is significant heterogeneity across these data sources arising from multiple coding systems, languages, structures, content and governance which complicate the implementation of a CDM across European data.”¹⁵ Comparing the FDA's Sentinel to the mosaic of European systems and considering new legislative developments, such as the GDPR, the EMA concludes that achieving the same level of reassurance without exerting as much control will be difficult.¹⁵

In the light of these challenges, how European initiatives to enhance CDM and RWD uses will evolve in the coming years is unclear. Still, the breadth, depth,

and quality of RWD can be expected to increase, and the update problem will obviously require the EMA, the European Commission, and the member states to figure out how to improve CDM and RWD uses to address it. This step will require concrete health sector-specific initiatives and approaches, such as more harmonised European health system infrastructures, that go beyond the laudable principles and ambitions proposed in the February 2020 Commission paper, communication, and report. Even overcoming these important challenges will not be enough, as systems and processes using the harmonised infrastructure and data would still need to be built in order to manage the update problem using, for example, appropriate monitoring and risk management practices.

The challenges of new trade-offs

To realise the full potential of AI, strategic decisions have to be made and competing interests and values have to be balanced, which necessarily come along with trade-offs—often relating to privacy, intellectual property rights, accountability, transparency, explainability, performance, bias, and discrimination.¹⁶ For example, the European Commission emphasises in its paper that it sees trustworthiness as a prerequisite for the uptake of AI.¹ For AI systems to be classified as trustworthy, according to the High-Level Expert Group on AI set up by the Commission, seven essential requirements need to be fulfilled: “(1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) societal and environmental well-being, and (7) accountability.”¹⁷

But one level deeper, many open questions exist. For example, if transparency includes the explainability of decisions made by AI systems,¹⁷ it probably involves trade-offs regarding cybersecurity, privacy, accuracy, and intellectual property protection and innovation. Similarly, accountability is complex and ill-defined for AI trained on large, interconnected data with often probabilistic algorithms and possibly open-source components. Concepts like fairness might mean different things across different countries and stakeholders; indeed, multiple AI fairness concepts exist with incompatibilities and trade-offs among them.¹⁸ What should be done if a more accurate algorithm has the effect of reinforcing existing deprivations of worse-off groups—is that a reason to move to a less accurate but more solidaristic one? What if the EU trade-offs eventually embedded in European AI products and services are not in line with those of other countries? Will that drag Europe behind in global AI competition or in new and unclear values-based AI trade wars, and if so, is that a necessary evil to standing on principle? The Commission paper has not scratched the surface of these difficult questions, and that fact will be essential for the next steps in AI governance.

The path ahead

There is still time for the Commission to refine its AI approach. As this iterated process continues, the Commission must deepen its engagement with the identified issues and further clarify how such a new EU regulatory framework could be implemented quickly and efficiently. Although the current white paper acknowledges the challenges of the update problem, plenty of new processes and rules need to be put in place to ensure appropriate monitoring of AI systems when these evolve—and possibly improve. The support and development of international standards for designing and monitoring AI systems, such as those developed by professional organisations like the Institute of Electrical and Electronics Engineers or the International Organization for Standardization, is a crucial component for regulatory success.¹² The development of regulatory frameworks that consider the complete lifecycle of AI systems, from establishing the quality of all data used for their initial development to managing risks of the systems becoming misaligned with the environments they operate in or becoming exposed to new cybersecurity attacks to their final decommission, might be necessary. As the FDA among others has recognised, this might not yield a one-size-fits-all approach and it might be necessary to assess the maturity of organisations developing AI systems (eg, following good machine learning practices) in determining the correct regulatory balance.¹⁹

An important characteristic of AI that the Commission needs to consider is AI’s ability to adapt to different users and contexts—much like recommender systems change their recommendations depending on the customer. It is therefore crucial to consider AI not as standalone products, but as complex socio-technical systems with many interacting components.¹³ For example, the behaviour of AI systems depends, among others, on (1) third party data (if any) used to train it, some of which might be open to algorithmic choices made by researchers and developers; (2) how users might provide data to the system while using it; (3) how algorithms (some of which might be probabilistic in nature and thus difficult to assess and fully control) are designed to adjust to or to possibly ignore some input data; (4) how people might make decisions given input from the AI—in particular, behavioural effects such as people becoming overconfident or changing their risk attitudes when using AI to make decisions. Without consideration of the entire system, these technologies might hurt rather than improve health-care quality.^{13,20}

The Commission might also need to better separate issues with data versus algorithms—a relevant but fine line to balance. For example, AI systems are the result of tuning possibly thousands of parameters of mathematical equations. The end product is determined, among others, by the data selected to train the algorithms, the specific algorithms used, and trade-offs made by the algorithm developers—eg, balancing accuracy and fairness or

explainability. Values might be encoded in any of these steps. For example, retraining non-EU AI systems with European data might negatively impact accuracy without necessarily improving fairness if some parameters determining trade-offs between these two are specifically hard coded in the algorithm itself. Algorithms might be fair even if the data used to train them capture human biases—a goal of fairness in machine learning innovations currently being developed.²¹ Finally, policy makers must consider processes for continuous updating of AI policies, guidelines, and regulations in order to flexibly and efficiently adapt to AI innovations some of which might also solve current challenges, such as so-called federated learning,^{22,23} multi-task learning,²⁴ or privacy-preserving machine learning,²⁵ or lead to new challenges.

Conclusion

The European Commission has done a good job in setting out a European approach for AI. However, implementing this approach in health care will be more difficult. It will require careful balancing of core values, detailed consideration of nuances of health and AI technologies, and a keen eye on the political winds and global competition.

Contributors

All authors contributed equally to this Viewpoint.

Declaration of interests

IGC, SG, and TM were supported by a grant from the Collaborative Research Program for Biomedical Innovation Law, a scientifically independent collaborative research program supported by a Novo Nordisk Foundation grant (NNF17SA0027784). IGC served as a bioethics consultant for Otsuka on their Abilify MyCite product. The company neither funded the preparation of this Viewpoint nor played a role in its drafting or review. SG received funding from the German Federal Ministry of Education and Research, from April 1, 2016, to March 31, 2018, outside the submitted work. TM reports grants from the University of Copenhagen and the Wallenberg Foundation, outside the submitted work; and a part-time advisor in legal matters role for X-officio. TE declares no conflict of interests.

References

- European Commission. White paper. On artificial intelligence—a European approach to excellence and trust. https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (accessed March 9, 2020).
- European Commission. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A European strategy for data. <https://op.europa.eu/en/publication-detail/-/publication/ac9cd214-53c6-11ea-aece-01aa75ed71a1/language-en> (accessed March 9, 2020).
- 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. https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf (accessed March 9, 2020).
- Von der Leyen U. A Union that strives for more. My agenda for Europe. https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf (accessed March 9, 2020).
- Gerards J. Margin of appreciation and incrementalism in the Case Law of the European Court of Human Rights. *Hum Rights Law Rev* 2018; **18**: 495–515.
- Prince AER, Schwarcz D. Proxy discrimination in the age of artificial intelligence and big data. *Iowa Law Rev* 2020; **105**: 1257.
- Wachter S. Affinity profiling and discrimination by association in online behavioural advertising. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3388639 (accessed April 28, 2020).
- Gerke S, Minssen T, Cohen IG. Ethical and legal challenges of artificial intelligence-driven health care. In: Bohr A, Memarzadeh K, eds. *Artificial intelligence in healthcare*, 1st edn. London: Elsevier, 2020.
- Price WN 2nd, Kaminski ME, Minssen T, Spector-Bagdady K. Shadow health records meet new data privacy laws. *Science* 2019; **363**: 448–50.
- Minssen T, Gerke S, Aboym M, Price N, Cohen IG. Regulatory responses to medical machine learning. *J Law Biosci* 2020; published online April 11. DOI:10.1093/jlb/lsaa002.
- Allen A. Lost in translation: epic goes to Denmark. <https://www.politico.com/story/2019/06/06/epic-denmark-health-1510223> (accessed March 9, 2020).
- Babic B, Gerke S, Evgeniou T, Cohen IG. Algorithms on regulatory lockdown in medicine. *Science* 2019; **366**: 1202–04.
- Gerke S, Babic B, Evgeniou T, Cohen IG. The need for a system view to regulate artificial intelligence/machine learning-based software as medical device. *npj Digital Med* 2020; **3**: 53.
- FDA. FDA's Sentinel Initiative. <https://www.fda.gov/safety/fdas-sentinel-initiative> (accessed April 28, 2020).
- EMA. A common data model for Europe? – Why? Which? How? https://www.ema.europa.eu/en/documents/report/common-data-model-europe-why-which-how-workshop-report_en.pdf (accessed April 28, 2020).
- Minssen T, Rajam N, Bogers M. Clinical trial data transparency and GDPR compliance: implications for data sharing and open innovation. *Sci Public Policy* 2020; published online March 4. DOI:10.1093/scipol/scaa014.
- High-Level Expert Group on AI. Ethics Guidelines for trustworthy AI. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> (accessed March 9, 2020).
- Kleinberg J, Mullainathan S, Raghavan M. Inherent trade-offs in the fair determination of risk scores. *arXiv* 2016; published online Nov 17. <https://arxiv.org/pdf/1609.05807v2.pdf> (preprint).
- FDA. Proposed regulatory framework for modifications to artificial intelligence/machine learning (AI/ML) - based software as a medical device (SaMD). <https://www.fda.gov/media/122535/download> (accessed April 28, 2020).
- Philpotts LE. Can computer-aided detection be detrimental to mammographic interpretation? *Radiology* 2009; **253**: 17–22.
- Barocas S, Hardt M, Narayanan A. Fairness in machine learning. <https://fairmlbook.org/pdf/fairmlbook.pdf> (accessed April 28, 2020).
- Konečný J, McMahan HB, Ramage D. Federated optimization: distributed machine learning for on-device intelligence. https://www.maths.ed.ac.uk/~prichat/papers/federated_optimization.pdf (accessed April 28, 2020).
- NVIDIA. Federated learning brings AI with privacy to hospitals. <https://healthcare-in-europe.com/en/news/federated-learning-brings-ai-with-privacy-to-hospitals.html> (accessed April 28, 2020).
- Evgeniou T, Michelli CA, Pontil M. Learning multiple tasks with kernel methods. *J Mach Learn Res* 2005; **6**: 615–37.
- Dwork C. Differential privacy. In: Bugliesi M, Preneel B, Sassone V, Wegener I, eds. *ICALP. 2006 Automata, languages and programming*. Heidelberg: Springer, 2006: 1–12.

Copyright © 2020 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.