This article has been accepted for publication in 2025 in LAW, INNOVATION AND TECHNOLOGY, published by Taylor & Francis.

The European Union's Artificial Intelligence Act and Trust:

Towards an AI Bill of Rights in Healthcare?

Barry Solaiman

Summary

The AI Act adopts a risk-based regulatory framework, categorizing AI systems according to the level of risk they present and imposing corresponding regulatory requirements. While the regulation aims to cultivate trust by focusing on risk mitigation, it has faced criticism for its insufficient attention to the ethical and relational factors that are crucial for establishing genuine trust, especially within healthcare settings. In healthcare, trust encompasses a complex network of relationships involving patients, providers, and regulatory bodies, and is influenced by personal interactions, transparency, ethical practices, and the perceived intentions behind AI technologies and their implementers. This paper argues that trust in AI could be better fostered in healthcare through a Bill of Rights for AI that explicitly incorporates trust-building measures into the regulatory framework. Such an initiative could clarify the role and expectations of AI in healthcare, making it more trustworthy for both patients and healthcare providers.

Keywords

Artificial Intelligence Act (AI Act), Trust, Trustworthiness, Bill of Rights, Healthcare

1. Introduction

Would European Union (EU) citizens trust they are being safeguarded from the risks of artificial intelligence (AI) where the law regulates the cumulative amount of compute used for AI training that is measured in floating point operations (FLOPs) greater than 10^25'? [1]

The disconnection between the language used in the Artificial Intelligence Act (AI-Act) and the patient on the ground is perhaps symbolized no better than in Article 52a(2) of the regulation above. That article is the culmination of a protracted process beginning formally in 2021 that evolved considerably during a period in which each revision to the law exponentially added to the complexity of the regulation. With the dust finally settling, what has emerged is a behemoth-like bureaucratic regulation that links the palatability of risk to trust. While the EU ought to be applauded for developing the most comprehensive law on AI in the world to date, this article argues that much more work is needed to enhance trust in the use of AI in healthcare, building on the foundations established by the regulation.

A lack of regulation and inadequate regulations might lead to widespread harm and, consequently, loss of trust in AI in health.[2] The lack of regulation is being addressed with the AI-Act. That Act forms only one part of an emerging regulatory framework under development, which includes the Artificial Intelligence Liability Directive (AILD),[41] a revised Product Liability Directive (PLD),[42] and the European Health Data Space (EHDS).[43] It is beyond the scope of this paper to appraise all these

mechanisms. Instead, the focus here is solely on the AI-Act, its inadequate approach to building trust in healthcare, and what should be done to remedy that gap.

Much attention has already been given to the intersection of AI and trust. Policymakers are concerned that society will either trust AI too little (and thus be unable to reap the benefits) or trust AI too much (and with it, the implications of AI errors causing harm).[3] It has been said that AI's successful adoption in society 'hinges on trust'.[4] Properly regulating AI may help deal with these concerns, but regulation alone will not directly establish trust.[3] The following sections examine what 'trust' means in healthcare, how the AI-Act incorporates trust as a concept, and whether its provisions enhance trust. For policymakers, it proposes that a Bill of Rights for AI in health tied to the Act could help to remedy the existing gaps in the regulation of strengthening trust in healthcare.

2. The Concept of Trust in Healthcare

The subject of trust in healthcare has long been examined.[5] Trust has been conceptualized differently, using terminology often interchangeably without a clear definition.[6] Consequently, the lens of analysis can become conceptually complex. In generic terms, trust can be defined as an expectation or attitude of 'the trusting agent that the trusted agent will act in a predetermined manner'.[3]

In healthcare, there must be trust between healthcare professionals and patients, doctors and their employers, between those parties and regulators, suppliers, developers, and so much more. If trust involves interpersonal relationships, such as between healthcare professionals and patients, then different rules governing those relationships will affect the level of trust between people. For example, fiduciary law uses restorative remedies for any unexpected consequences of relationships based on trust.[7] The relationship between patients and healthcare professionals requires honesty and openness to foster trust in those professionals and the health system more broadly.[8] Another example is the trust between clinicians and their employer healthcare organizations.[9] Events can impact trust in that relationship, which occurred during COVID-19 when the increased demands on professionals led to moral injury and feelings of betrayal by leaders.[9]

The positive and negative effects on trust depend on what different people value. At a micro-relational level, studies examining trust between doctors and patients find that trust is positively affected by the doctor's behaviour, the perceived comfort levels of the patient with the doctor, the personal involvement with the patient, and, to a lesser extent, the doctor's cultural competence and physical appearance.[10] Some patients value comfort, so trust is enhanced when doctors exhibit more comforting and caring behavior.[9] For others, trust is influenced more by having a personally involved clinician who offers personalized care.[10] Some patients care that the doctor shares common traits like language, religion, cultural beliefs, and values for emotional connectedness.[10] The doctor's behaviour is also critical. This involves their 'communication skills, smiling face, kindness and non-discrimination'.[10] Communicating the truth is essential, but that communication should be patient-centred and patient-friendly.[10,11]

While pinpointing individual factors is context-dependent, some fundamental principles thread through healthcare – principles that, if undermined, would be detrimental to trust in the medical field. These are examined in more detail later in this paper, but one example is informed consent, which is 'required because of trust and without regard to the effects on trust'.[12] Patients place a special kind of trust in doctors who have special responsibilities to the patient, such as sharing honest and clear information with them so they can make informed decisions about their care.[12] This is called a 'predicated' stance toward trust, whereby ethics and law impose particular obligations because trust is a factual premise, so 'trust is the source of the obligation, not its object'.[12] Confidentiality is another example. Patients must trust that the information they share with their doctor is kept private and secure, and laws require this.[12] This has been called a 'supportive' stance towards trust, whereby rules are created to maintain and promote trust.[12] Where it is perceived that trust does not exist or cannot be sustained, this may be used to justify a regime to institutionalize trust.[12]

Trust can also operate at a macro level. People value health systems not only for the individual care they receive but also for their broader contribution to the well-being of society.[13] It has been argued that it is important to develop the legitimacy of state action within health systems to garner trust.[13] The state has a central role, and its actions must be seen as legitimate in persuading patients to cooperate and accept health interventions.[13] Trust must be built in the state and its agencies to establish legitimacy. Even in this context, building trust involves examining personal behaviours between healthcare providers, healthcare professionals and patients, between healthcare professionals and their employers, and between the public and private sectors.[13] Managerial and organisational practices can help to foster trust. For example, by creating spaces for caring engagement, dialogue, and interpersonal interactions.[13]

These considerations about trust should be borne in mind when examining the EU's AI-Act, which will ultimately impact healthcare. However, the EU is not a state, and does not have competence to regulate healthcare systems. Instead, the EU regulates the internal market, which may lead to some interface with healthcare systems at a state level. This paradigm underscores a limitation of the supposed trust-building aims of the AI-Act and how the principles of trust manifest in healthcare and the state level.

3. Trust and the AI-Act

The AI-Act is not explicitly designed for healthcare or other sectors. Instead, it aims to 'promote the uptake of human centric and trustworthy artificial intelligence while ensuring a high level of protection of health, safety, fundamental rights'.[1] This is to be achieved by mitigating risks inherent in AI through the Act's risk-based regulatory approach that imposes restrictions and controls on AI deemed risky. While the emphasis on trust seems clear, the term 'trust' seldom appears in the 245-page regulation (appearing only once in the articles and a dozen times in the recitals with a simple keyword search of the regulation).

According to the recitals, trust will be achieved through a 'consistent and high level of protection' throughout the EU.[1] The concern is that AI will significantly impact society, so there is a need to build

trust (aligning with the macro-level considerations about trust noted above). The regulation offers a method for 'building' trust - through a regulatory framework developed 'according to Union values enshrined in Article 2 TEU, the fundamental rights and freedoms enshrined in the Treaties, the Charter'.[1] AI should be 'human-centric' and serve as a 'tool for people, with the ultimate aim of increasing human well-being'.[1]

3.1 The Basis of Trust

Given the work that preceded it, the articulation of trust in the AI-Act is unsurprising. In 2019, the European Commission appointed the High-Level Expert Group on AI (AI HLEG) to develop Ethics Guidelines for Trustworthy AI.[1,14] The Guidelines establish a framework for trustworthy AI, how to realise such trustworthy AI systems, and how they can be assessed for their trustworthiness.[14] Such ethical guidelines have been criticised for being 'useless'.[48] Indeed, analysing the basis for the Guidelines reveals several shortcomings.

'Trustworthy' AI should have three components according to the Guidelines. First, it should comply with all applicable laws and regulations (primary EU law, like the treaties and the Charter of Fundamental Rights, and secondary law, such as the GDPR, the ECHR, and medical device regulations).[14] Second, AI systems should adhere to ethical principles and values, such as respect for human autonomy, prevention of harm, fairness, and explicability.[14] Third, AI should be robust from a technical and social perspective.[14] This includes standardization and certification (such as ISO standards), which can play an important role in garnering trust.[14] All components should work together and overlap in their operation, and act as a check on one another.[14]

In 2020, the European Commission stated that 'trustworthiness is a prerequisite' for the uptake of AI, tying the concept of trust to values and fundamental rights, such as ensuring human dignity and protecting privacy.[15] It recommended the development of an 'ecosystem of trust'.[15] The same year, the Parliamentary Assembly of the Council of Europe stated that there was a 'big question of trust regarding some AI applications in health'.[16] Regarding the COVID-19 pandemic, it noted that 'had there been a trusted and well-defined regulatory framework, maybe AI could have had a much larger positive impact on the managing of this pandemic'.[16]

The Guidelines form part of a theoretical and governance framework concerning trust in AI. The AI-Act cites seven non-binding principles from those guidelines that supposedly help ensure AI is trustworthy and ethically sound. Namely, human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being, and accountability.[1] Together, these principles are the basis of the conception of trust in the regulation. However, that basis is not captured in the resulting law.

3.2 Equating Risk and Trust

The AI-Act does not meaningfully entrench the concepts designed to enhance trust. Instead, it creates a risk-based approach towards regulating AI vaguely linked to trust-building.

The recitals state that the regulation applies mandatory requirements to high-risk AI systems to mitigate 'risks' from placing them on the market and to 'ensure a high level of trustworthiness'.[1] To ensure a 'high level of trustworthiness,' a conformity assessment will be required.[1] This is the final recital that mentions trust explicitly. After that, the only actual article in the regulation that mentions 'trust' is Article 1(1) to reiterate the articulation in the recitals that the regulation exists to promote the 'uptake of human centric and trustworthy artificial intelligence'.[1] Readers are left to ponder how the other provisions may enhance trust in congruence with the stated ethical basis. That approach does not inspire confidence that trust-building has been conceptualized coherently nor implemented systematically as an underlying aim of the regulation.

Instead, the provisions distinguish between AI systems that pose unacceptable risks, high risks, and low risks.[1] Devices posing an unacceptable risk, such as systems that use social scoring or manipulative AI, are prohibited. Low risk systems are not captured by the law, encompassing apps one might find on an app store, such as wellness apps. For healthcare, the regulation is most likely to capture systems subject to the EU's Medical Device Regulation (MDR) and the regulation on in-vitro diagnostic medical devices (IVD).[17,18] Those regulations are noted in the AI-Act's Annex, meaning that medical devices that incorporate AI will be subject to additional rules. For example, a medical device provider must follow the MDR's conformity assessment, and the AI-Act's conformity assessment will be an additional element within that assessment process.

It is doubtful that additional conformity assessments on 'risk' would do much to enhance trust in AI. Devices for the MDR should bear 'CE marking' (Conformité Européenne) to demonstrate conformity with the EU's safety requirements.[17] However, information submitted for CE marking is confidential, and does not indicate device effectiveness, only that it is legally compliant with minimum EU-level safety standards.[19] CE marking is based on 'unsubstantiated trust' that evidence of safety and effectiveness will be provided later once the product is on the market.[19] Consequently, it has been argued that the MDR is not aligned with international ethical standards for clinical research.[19] It must be queried whether stakeholders will trust AI devices brought to market on already flawed regulatory mechanisms.

Further, the AI-Act enshrines factors for determining high-risk classifications, but it is unclear what the factors mean as a practical matter for compliance.[40] Under art 7(2), the intended purpose of the system is considered; whether special categories of data are processed (like health data); whether humans can override autonomous systems; the extent to which a system has already caused harm to health and human safety; the extent of imbalances of power, such as users being in vulnerable positions; and how easy it is to reverse the decision of the AI system.[1] These are all important considerations, but the regulation does not clearly tie these requirements to its trust-building beyond creating additional requirements on top of existing conformity assessment requirements from other regulations.

Indeed, it has been argued that the AI-Act's understanding of trust in terms of acceptability of risk is a 'simplistic conceptualization of trust', and the EU oversells its regulatory ambition.[20] While there is a relationship between trust, trustworthiness and acceptability of risks, the AI-Act 'conflates' the

distinction between what experts deem acceptable risks, and whether society trusts the AI systems eventually implemented.[20] Trust is a longitudinal process of 'controls, communication, and accountability', rather than a binary analysis of AI risk.[20] What fosters public trust is a domain-specific question that 'casts doubt on the effectiveness of a horizontal regulatory law such as the AI Act'.[20]

The AI-Act also creates AI auditors (notified bodies) to verify the conformity of high-risk systems, but there are concerns that those bodies may be influenced by private industry.[20] While notified bodies must be independent of the provider of the high-risk AI system, they charge a fee for their services, creating the obvious incentive to receive repeat commissions from AI developers.[20] Lay people who cannot assess the trustworthiness of an AI system will need to trust these accountability mechanisms.[20] It is doubtful that misaligned incentives embedded within the regulatory framework will do much to garner trust.

These criticisms are underscored when examining the AI-Act's approach to generative AI (GenAI), like ChatGPT. The use of ChatGPT raises significant concerns about trust because there is a very worrying trend of some healthcare professionals utilizing it in clinical practice despite it not being designed or approved for medical use, potentially putting patients at substantial risk.[1] Provisions cover GenAI posing a 'systemic risk,' meaning where the floating point operations (FLOPs) are greater than 10^25.[1] In other words, GenAI with capabilities of GPT4 or higher. Systems captured by the regulation must comply with additional transparency requirements. The practical consequence is that patients must be informed they are interacting with a GenAI system if it is not 'obvious' to them.[1]

The technicality of the 'flops' terminology used in the regulation, with its attendant focus on 'system risk', belies the resulting protections for users on the ground. Even if healthcare professionals notify their patients when using ChatGPT (satisfying the AI-Act), they should not use it in the first instance because its use may fall below the legal standard of care. It is also worth noting that (at the time of writing) GPT4 is an additional paid service, whereas the free version runs on GPT 3.5 (not captured by the AI-Act at all). There is a significant chasm of standards that are pertinent to the healthcare domain and a disconnect between the basis of trust underlying the Act and the safety standards needed for stakeholders in healthcare.

4. Discussion: Refocusing on Trust Through an AI Bill of Rights for Healthcare

The AI-Act falls short of connecting its premise on trust with its risk-based approach. However, it does create space for further developments that stakeholders should build on to enhance trust in the health domain. Unfortunately, the space created by the AI-Act encourages the 'development of voluntary best practices and standards' that take into account ethical principles.[1] Voluntary codes are hardly inspiring because they could inevitably lead to incongruent interpretation and application, and the scope of the codes proposed in the AI-Act is minimal. However, that is the space created by the AI-Act which will likely be built upon. This may be an area where some 'public benefit' can be established as called for elsewhere.[49] Here, it is recommended that a voluntary code can take the form of an AI Bill of Rights for Health to accompany the Act. The articulation of a 'bill of rights' is proposed to connect trust-building principles directly to the aims of the AI-Act on trust. Also, it will be easier

to educate patients about the protections they can expect for using AI in health when framed as rights they can expect when compared to technical requirements about conformity. The proposal here is intended to encourage debate and is not exhaustive.

4.1 Systemic Considerations

It is first helpful to note the distinction between the 'trust' and 'trustworthiness' concepts because they indicate the scope of a potential Bill of Rights. Trust and trustworthiness are separate concepts; the former does not necessarily lead to the latter.[6] Trust encapsulates the idea of placing trust in someone or something. Trust is a prerequisite to trustworthiness – one must trust that something is trustworthy to have trust in it.[21]

Unever et al., note that this idea of trust portrays the approach taken in the EU's Guidelines for Trustworthy AI that examine trust from the perspective of both the inherent properties of the AI system as well as the qualities of the broader socio-technical system in which AI exists.[21] The authors emphasize that trust must derive from the operation of the entire socio-technical ecosystem within which AI exists as one component. Trust in AI as a standalone device is a concept that can be disproved because liability systems are premised on the responsibility of the clinician or their employer.[21] They argue that AI itself cannot have the capacity to be trusted because it does not have an emotive state and cannot be held responsible for its actions.[21] Nevertheless, while AI lacks morality, those developing AI have moral agency.[22] Attention should, therefore, be given to factors that affect trust within the overall healthcare system, such as ethical principles of transparency, responsibility, accountability, fairness, and privacy.[21]

AI requires trust in socio-technical systems, which involves the interconnection of different relationships and actors, such as healthcare institutions, healthcare professionals, patients and their families, AI developers, and more.[7] Regulation requires a multi-disciplinary and multi-stakeholder strategy that involving AI creators, the users of those systems, and politicians.[23] This holistic approach is essential because concerns may spill over from one domain to the other. If people do not trust the government to protect their data in one sector, they may not trust the same government to protect their data in other sectors like healthcare.[6] Trust, therefore, operates on a multidimensional level, requiring different values to be responsive to patient safety to ensure physical, social, and emotional well-being.[21] The responsibilities of those interacting with AI and the conditions attached to their use of AI within that system should be clear.[21]

Similar calls made by others for trust in AI to encompass the 'entire relationship-building process and not on the trait of trustworthiness alone'.[6] However, this is particularly complex because AI is used in different healthcare contexts and environments and integrated in different ways. Ensuring trust in different clinical contexts requires institutions to enable trust from the beginning. For example, from training AI to the eventual day-to-day usage in clinical practice, with ethical safeguards ensuring accountability for possible harms.[7] This requires designating responsibilities for different staff, such as doctors and clinicians, and outlining how the safeguards govern workflows concerning diagnosis and treatment using AI tools.[7] The use of AI in these contexts entails a holistic concept whereby AI,

as a subsystem operating in a more extensive healthcare system, should serve the overall trust in that system.[7]

The problem is that these trust requirements move beyond the AI-Act's scope, indicating the need for a broader regulatory framework. The AI-Act is premised on allowing high-risk systems to market and does not cover the responsibilities of hospitals, clinicians, and their workflows. Another problem is that people are less likely to trust AI than a human doctor, even where AI provides the same level of care as a doctor.[24] People perceive that AI cares less than a human doctor and cannot share similar values. Patients do not trust diagnoses provided by AI as much as those provided by human healthcare professionals.[25] There is a 'substantial resistance to the use of AI' in diverse populations of patients.[25] Patients' concerns about perceived care and value are 'significant determinants of trust in risk perception research.'[24] This leads to decreased trust in AI.

It is unclear whether the AI-Act may interact with the fundamental question of the doctor-patient relationship. Nevertheless, the Act must promote the values necessary for trust in AI in healthcare so that any product coming to market is checked for compliance with such values at the front end. These values have been examined for several years and are now encapsulated within health, AI, the law, and bioethics.[26] Indeed, there is a positive relationship between AI trust and the ethical governance of AI.[23] The values are typically premised on consent, medical liability, data accuracy, privacy, bias, security, efficacy, safety, and transparency.[22,27,28]

4.2 Values for Trust

These values cover longstanding recommendations in academic literature on ensuring provider competence, protecting the patient's interests, and ensuring information integrity.[29] Below are some non-exhaustive examples of what could be included in a Bill of Rights under the AI-Act to promote the values.

One primary concern for doctors and patients is informed consent and explainability.[30] Patients should be fully informed about their care to make informed decisions. AI is problematic because the reasons for an AI's decision or recommendation are often unknown or unknowable owing to the complex neural network underpinning its operation. The opaqueness of AI systems can make attaining transparency, accountability, and responsibility difficult.[6] This is called the black box problem, and one aim is to enhance the 'explainability' of AI decisions so that people can understand the underlying logic. However, explainability in AI is easier said than done. One can disclose the learning algorithms, but they can lack data features and detailed explanations of the weights that influence the eventual output, which may do little to explain the underlying logic of a decision.[31] A Bill of Rights should require meaningful disclosure of information that experts can access to alleviate matters of broad public concern in health.[31] Disclosure could enable experts to determine risks for informed consent that a particular algorithm poses in terms of explainability. That is not to say that explainability is a panacea. Some methods offer post hoc rationalisations of a black box, which are unlikely to assist in understanding the inner workings of the AI system.[32] Nevertheless, the AI-Act's framework should work towards enshrining the right of patients to obtain sufficient information for their care.

Data considerations present other critical facets of trust. AI requires large representative datasets, and patients may benefit from AI using such data. However, patients do not trust that governance systems will keep their data confidential.[2] The problems of deidentification, reidentification, and inferences have been well documented.[33] A Bill of Rights might require compliance with specific provisions in the General Data Protection Regulation (GDPR) concerning the use of sensitive data or other stricter standards of confidentiality found in health. Another problem is that the data used to train systems may be biased, producing incorrect recommendations for the demographic it is used on.[44] A Bill of Rights can require additional checks and verification by creators of AI systems, anyone operating the system, and anyone interpreting its outputs to check for biases.[23] Those checks could help ensure that diverse and representative data was used in AI development.[34] The European Health Data Space (EHDS) may also address this challenge. It will allow researchers and others 'non-discriminatory access to health data and the training of artificial intelligence algorithms'.[1]

Accountability for AI decisions is also essential. Patients must trust both the healthcare professional's decision (made with the assistance of AI) and the AI system itself.[22] If a doctor relies on an AI's recommendation, harm may result to patients, raising questions of whether the doctor, hospital or AI developer is liable, particularly under tort law.[36] The AI-Act does not cover liability concerns, but the proposed Artificial Intelligence Liability Directive (AILD) and the revised Product Liability Directive (revised PLD) may. A Bill of Rights associated with the AI-Act should enshrine the right to redress for harms through existing standards of care in tort law and include forthcoming directives. However, it should move beyond the right to legal redress by requiring a human point of reference to answer questions and resolve problems where harms occur.[23] This could mean a responsible person or standing committee within the hospital setting convened to examine AI incidents.[45]

A final example is safety and effectiveness.[35] Patients should expect that AI outputs are accurate and safe for the context in which they are applied. Datasets should be reliable and valid, but they often are not because of a lack of training data.[46,47] For example, developing accurate AI systems for psychiatric wards is challenging because of the small sample sizes of data available.[37,38] Practical limitations can also affect accuracy. Monitoring systems used in clinical contexts may require multiple sensors with overlapping fields of view to track patients in real-time and make accurate predictions.[39] A Bill of Rights should require that systems are designed to be accurate to help ensure they are safe and effective.

5. Conclusion: Moving Forward

These are some areas where a Bill of Rights could stipulate standards for using AI in healthcare that instil trust. Rather than focusing purely on risk classification and equating that classification to enhancing trustworthiness, measures should be set out that tie trust specifically to health. In the first instance, the Bill of Rights could operate as a voluntary code within the regulation's remit. Provisions focussing on data, consent, and safety will align with the aims of the Act and be directed towards those most affected by AI – the end users and the patients subject to their decisions.

References

1. AI-Act, European Parliament. Provisional Agreement Resulting from Interinstitutional Negotiations, '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' (AI-Act) COM (2021) 0106 (COM), Final Draft version as of 21 January–2 February 2024
2. Da Silva M, Horsley T, Singh D, et al. Legal concerns in health-related artificial intelligence: a scoping review protocol. Syst Rev. 2022;11:123. doi:10.1186/s13643-022-01939-y.
3. Tamò-Larrieux A, Guitton C, Mayer S, Lutz C. Regulating for trust: Can law establish trust in artificial intelligence?. Regulation and Governance. 2023;1:2.
4. Erdélyi OJ, Goldsmith J. Regulating artificial intelligence: Proposal for a global solution. Government Information Quarterly. 2022;39(4):10
5. Hall MA. Law, Medicine, and Trust. Stanford Law Rev. 2002;55(2):463–527. Available from: https://doi.org/10.2307/1229596
6. Gille F, Smith S, Mays N. Why public trust in health care systems matters and deserves greater research attention. J Health Serv Res Policy. 2015;20(1):62-64. doi:10.1177/1355819614543161
7. Unver MB. Governing fiduciary relationships or building up a governance model for trust in AI? Review of healthcare as a socio-technical system. Int Rev Law Comput Technol. 2023;37(2):198-226. doi:10.1080/13600869.2023.2192569
8. Quick O. Duties of Candour in Healthcare: The Truth, the Whole Truth, and Nothing but the Truth? Med Law Rev. 2022 Spring;30(2):324–347. Available from: https://doi.org/10.1093/medlaw/fwac004
9. Baron RJ, Simpson L. Understanding Clinician Trust in Health Care Organizations Should Be a Research Priority. JAMA Health Forum. 2023;4(5):e231015. doi:10.1001/jamahealthforum.2023.1015
10. Gopichandran V, Chetlapalli SK. Factors influencing trust in doctors: a community segmentation strategy for quality improvement in healthcare. BMJ Open. 2013;3:e004115. doi:10.1136/bmjopen-2013-004115
11. Thom DH. Physician behaviors that predict patient trust. J Fam Pract. 2001;50:323–8.
12. HALL MA. The Importance of Trust for Ethics, Law, and Public Policy. Camb Q Healthc Ethics. 2005;14(2):156-167. doi:10.1017/S096318010505019X
13. Gilson L. Trust and the development of health care as a social institution. Soc Sci Med. 2003;56(7):1453–1468.
14. European Commission, Independent High-Level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy AI. 8 April 2019. Available from: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
15. European Commission. White Paper on Artificial Intelligence: a European approach to excellence and trust. Brussels, 19 February 2020. COM(2020) 65 final.
16. Committee on Social Affairs, Health and Sustainable Development. Artificial intelligence in health care: medical, legal and ethical challenges ahead. Rapporteur: Ms Selin Sayek Böke, Turkey, SOC. Parliamentary Assembly, Council of Europe, 22 Sept 2022.
17. (MDR) Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on Medical Devices, Amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and Repealing Council Directives 90/385/EEC and 93/42/EEC [2017] OJ L117/1.
18. (IVD) Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on In Vitro Diagnostic Medical Devices and Repealing Directive 98/79/EC and Commission Decision 2010/227/EU [2017] OJ L117.
19. Hulstaert F, Pouppez C, Primus-de Jong C, et al. Gaps in the evidence underpinning high-risk medical devices in Europe at market entry, and potential solutions. Orphanet J Rare Dis. 2023;18:212. Available from: https://doi.org/10.1186/s13023-023-02801-7
20. Laux J, Wachter S, Mittelstadt B. Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk. Regulation Governance. 2024;18(1):3-32. Available from: https://doi.org/10.1111/rego.12512

21. Unver MB, Asan O. Role of Trust in AI-Driven Healthcare Systems: Discussion from the Perspective of Patient Safety. Proc Int Symp Hum Factors Ergon Health Care. 2022;11(1):129-134. doi:10.1177/2327857922111026
22. Jermutus E, Kneale D, Thomas J, Michie S. Influences on User Trust in Healthcare Artificial Intelligence: A Systematic Review [version 1; peer review: 4 approved with reservations]. Wellcome Open Res. 2022;7:65. Available from: https://doi.org/10.12688/wellcomeopenres.17550.1
23. Omrani N, Rivieccio G, Fiore U, Schiavone F, Agreda SG. To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. Technological Forecasting & Social Change. 2022;181:121763. Available from: https://doi.org/10.1016/j.techfore.2022.121763.
24. Yokoi R, Eguchi Y, Fujita T, Nakayachi K. Artificial Intelligence Is Trusted Less than a Doctor in Medical Treatment Decisions: Influence of Perceived Care and Value Similarity. International Journal of Human–Computer Interaction. 2021;37(10):981-990. DOI: 10.1080/10447318.2020.1.
25. Robertson C, Woods A, Bergstrand K, Findley J, Balser C, Slepian MJ. Diverse patients' attitudes towards Artificial Intelligence (AI) in diagnosis. PLoS Digital Health. May 19, 2023. Available from: https://doi.org/10.1371/journal.pdig.0000237.
26. Solaiman B, Cohen IG. A Framework for Health, AI and the Law. In: Solaiman B, Cohen IG, editors. Research Handbook on Health, AI and the Law. Edward Elgar; 2024.
27. UK Department of Health & Social Care. Guidance: A guide to good practice for digital and data-driven health technologies. Updated January 19, 2021. Available from: https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology.
28. Mittelstadt B. The impact of artificial intelligence on the doctor-patient relationship. Council of Europe; December 2021. Available from: https://www.coe.int/en/web/bioethics/report-impact-of-ai-on-the-doctor-patient-relationship#:~:text=The%20potential%20human%20rights%20impact,account%20of%20well-being%3B%20.
29. Feldman R, Aldana E, Stein K. Artificial Intelligence in the Health care Space: How We Can Trust What We Cannot Know. Stanford Law & Policy Review. 2019;30:399. Available from: https://repository.uchastings.edu/faculty_scholarship/1753. Pages 413-416.
30. Solaiman B, Bloom M. AI, Explainability and Safeguarding Patient Safety in Europe: Toward a Science-Focused Regulatory Model. In: Cohen IG, et al., editors. The Future of Medical Device Regulation: Innovation and Protection. Cambridge University Press; 2022.
31. Dwork C, Minow M. Distrust of Artificial Intelligence: Sources & Responses from Computer Science & Law. Daedalus. 2022;151(2):309–321. DOI: https://doi.org/10.1162/daed_a_01918.
32. Babic B, Gerke S, Evgeniou T, Cohen IG. Beware explanations from AI in health care. Science. 2021 Jul 16;373(6552):284-286. DOI: 10.1126/science.abg1834. PMID: 34437144.
33. Price N. Problematic Interactions Between AI and Health Privacy. 2021 ULR 925 (2021). DOI: 10.26054/0d-th4e-sgvq.
34. Ueda D, Kakinuma T, Fujita S, et al. Fairness of artificial intelligence in healthcare: review and recommendations. Jpn J Radiol. 2024;42:3–15. DOI: https://doi.org/10.1007/s11604-023-01474-3.
35. Gerke S, Minssen T, Cohen G. Ethical and legal challenges of artificial intelligence-driven healthcare. In: Artificial Intelligence in Healthcare. 2020:295–336. DOI: 10.1016/B978-0-12-818438-7.00012-5. Epub 2020 Jun 26. PMCID: PMC7332220.
36. Price WN II, Gerke S, Cohen IG. Liability for Use of Artificial Intelligence in Medicine. In: Solaiman B, Cohen IG, eds. Research Handbook on Health, AI and the Law. Edward Elgar Publishing; 2024.

37. Solaiman B, Malik A, Ghuloum S. Monitoring Mental Health: Legal and Ethical Considerations of Using Artificial Intelligence in Psychiatric Wards. Am J Law Med. 2023 Jul;49(2-3):250-266. DOI: 10.1017/amj.2023.30. Epub 2024 Feb 12. PMID: 38344795.
38. Wells C, et al. Artificial Intelligence and Machine Learning in Mental Health Services: An Environmental Scan. CADTH Health Tech Rev. 2021;(1):1-23. Available from: https://www.cadth.ca/sites/default/files/attachments/2021-06/artificial_intelligence_and_machine_learning_in_mental_health_services_environmental_scan.pdf.
39. Haque A, et al. Towards Vision-Based Smart Hospitals: A System for Tracking and Monitoring Hand Hygiene Compliance. JMLR. 2017;68:75-77.
40. Minssen T, Solaiman B, Köttering L, Wested J, Malik A. Governing AI in the European Union: Emerging Infrastructures and Regulatory Ecosystems in Health. In: Solaiman B, Cohen IG, eds. Research Handbook on Health, AI and the Law. Edward Elgar; 2024.
41. (AILD) Commission, 'Proposal for a Directive of the European Parliament and of the Council on Adapting Non-contractual Civil Liability Rules to Artificial Intelligence (AI Liability Directive)' COM (2022) 496 final.
42. (PLD) Council Directive 85/374/EEC of 25 July 1985 on the Approximation of the Laws, Regulations and Administrative Provisions of the Member States Concerning Liability for Defective Products [1985] OJ L210.
43. (EHDS) Commission, 'Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space' COM (2022) 197 final 1.
44. Abràmoff MD, Tarver ME, Loyo-Berrios N, Trujillo S, Char D, Obermeyer Z, Eydelman MB; Foundational Principles of Ophthalmic Imaging and Algorithmic Interpretation Working Group of the Collaborative Community for Ophthalmic Imaging Foundation, Washington, D.C.; Maisel WH. Considerations for addressing bias in artificial intelligence for health equity. NPJ Digit Med. 2023 Sep 12;6(1):170. DOI: 10.1038/s41746-023-00913-9. PMID: 37700029; PMCID: PMC10497548.
45. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc. 2020 Mar 1;27(3):491-497. DOI: 10.1093/jamia/ocz192. PMID: 31682262; PMCID: PMC7647243.
46. Smith AA, Li R, Tse ZTH. Reshaping healthcare with wearable biosensors. Sci Rep. 2023;13:4998. DOI: 10.1038/s41598-022-26951-z.
47. Xue Y. A review on intelligent wearables: Uses and risks. Hum Behav & Emerg Tech. 2019;1:287–294. DOI: 10.1002/hbe2.173.
48. Munn L. The uselessness of AI ethics. AI Ethics. 2023;3:869–877. doi:10.1007/s43681-022-00209-w.
49. Prainsack B, Forgó N. New AI regulation in the EU seeks to reduce risk without assessing public benefit. Nat Med 2024. doi:10.1038/s41591-024-02874-2.