Carleton University

Individual Research Paper

EU AI Act as a gold standard for AI governance in health

Emmanuelle Rhéaume

INAF5706 Global Health Policy

Valerie Percival

November 27, 2025

Table of Contents

Introduction

Since the late 2010s, artificial intelligence (AI) has seen rapid and transformative growth. It has been reshaping how societies function and how governments deliver public services as it has the power to transform major areas like healthcare, economy, or security. It can, for instance, increase efficiency in public administration, optimize service delivery, and find solutions to challenges like climate change and public safety. However, many risks are associated with the expansion of such technology. Fundamental rights are endangered as AI can impact negatively employment, ethical values, privacy, intellectual property, and equality. Its exponential progress is what makes it so risky as policy and research cannot progress at an even pace. What drives that rapid progress can be explained by five factors: advances in computing power, the availability of large datasets, increased investments from private and public sectors, expanding research collaboration across institutions, and the growing global demand for AI. Researchers also agree that the launch of ChatGPT in 2022 marked a turning point in AI history, as it made AI so popular, that its development outpaced regulatory and ethical frameworks by far. (Cancela-Outeda, 2024)

In healthcare, AI's potential is especially interesting. As the healthcare sector continues to face major challenges, such as a shrinking workforce, the rise of chronic diseases, and the growing impact of climate change, AI is increasingly seen as a potential tool to help build stronger, more resilient health systems. AI systems are starting to be integrated into the delivery of care, and their benefits include enhanced diagnostic accuracy, reduced administrative burdens, and improved treatment outcomes. However, the growing role of AI in health raises even bigger ethical and governance questions as it is a critical sector and the technology is not regulated enough yet. The "black box" nature of many algorithms poses challenges. Indeed, developers cannot fully explain how decisions are made inside these systems because of their complexity. It therefore poses

challenges for accountability, safety, and trust. Moreover, the risks of biased algorithms, data leaks, unequal access, and misuse of health information show how important it is to have solid rules in AI governance in the health sector. (Cancela-Outeda, 2024)

Over the past few years, we have seen an evolution of the regulatory frameworks for AI. In the beginning, these frameworks mainly relied on ethical guidelines, voluntary principles and self-regulation. They had limited impact as they lacked enforcement mechanisms and relied on goodwill. (Birksedt et al., 2023; Cancela-Outeda, 2024) As AI application became more complex and high stakes, the need for binding regulation became urgent. In response, the European Union (EU) Artificial Intelligence Act (AIA) emerged as the world's first comprehensive AI law. This framework is designed to respond to the gaps identified in the regulations of AI globally. Firstly, it makes governance actionable with concrete requirements for high-risk AI systems. Secondly, it is enforceable since it has monitoring bodies and penalties for non-compliance. Thirdly, it operationalizes ethical principles like transparency, accountability, fairness, and safety into technical and procedural obligations. Finally, it provides structure and clarity. This major legislative development is significant not only because it regulates AI across sectors and address major gaps in AI governance, but also because it aims to set global norms through what researchers call the "Brussels effect". It could, indeed, influence developers outside of the EU. AI transcends borders and developers could want to adapt their products to match EU standards in the AIA. (Cancela-Outeda, 2024; van Kolfschooten & van Oirschot, 2024) This could result in influencing other regions into how they govern AI. However, there are still little evidence and understanding around AI governance implementation. (Birkstedt et al., 2023). The EU AIA is still very recent and whether or not it can be used as gold standard for AI governance in health remains to be proven. Moreover, questions remain about whether it can serve as a model for low-income countries (LICs)

and middle-income countries (LMICs), which face very different political, economic, and institutional contexts when building their own AI governance frameworks.

This research paper will explore whether the EU AI Act is a gold standard for AI governance in health, and whether LICs and LMICs can learn from it and use it as a benchmark for developing their own regulatory frameworks. To answer these questions, the paper will define what constitutes good AI governance in health, review existing AI regulatory frameworks, analyze the EU AI Act to identify both its innovations and its limitations, and evaluate which elements could be applicable to the contexts of LICs and LMICs.

This paper argues that the EU AI Act does set important standards and introduce valuable safeguards for AI in healthcare. However, it cannot be fully applied in LICs and LMICs without proper adaptation. LICs and LMICs can learn from the EU framework's risk-based approach, and transparency and governance principles, but must adapt its implementation to their local capacities, priorities, and constraints. Basic regulatory capacity, accessible data and data quality, privacy and security, biases and inequalities, human oversight, and public trust and community engagement are elements that were identified by the United Nations Development Programme (UNDP) as essential to be address in AI governance frameworks in health that take into account the socio-economic context of these countries. While the Act does address many of these elements, it might still be too technical, resource-intensive, and dependent on strong institutions. These findings make selective adaptation and context-sensitive implementation essential for LICs and LMICs to benefit from its principles without creating additional risks or inequalities.

Defining AI governance

AI and AI governance are terms that are still being defined. They do not have one universal definition. The Organisation for Economic Co-operation and Development (OECD) AI principles, describe AI system as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment". (OECD, 2024) AI governance is defined by Birksedt et al. (2023) as "a system of rules, practices, processes, and technological tools that are employed to ensure an organization's strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI followed by the organization". However, it is still an emerging research topic with only a few explicit definitions.

Existing work also mentions that AI governance is about finding the right balance between creating value and managing risks (Al Zadjali, 2020). It can involve building ethical and legal principles directly into AI systems, for example through value alignment or clearly defining goals (Aliman & Kester, 2019), and combining legal, ethical, and technical approaches to make AI more trustworthy (Cath, 2018). Other researchers highlight the practical tools used in governance. For example, Butcher and Beridze (2019) talk about ethics frameworks, transparency tools, and regulatory interventions, while Perry and Uuk (2019) talk about the broader societal, political, and economic implications of governing AI systems.

Research approach

This report is a qualitative literature review on articles gathered through two main platforms, Google Scholar and the Carleton Omni Library. It used key search terms related to the phenomenon

studied to navigate through all the literature available. To refine the research even further, the articles reviewed were the ones that were the most cited and peer-reviewed even though some are not peer-reviewed as the timeframe of this research is still very recent. Finally, the reference lists of the reviewed articles were also consulted to find relevant sources, and google searches were conducted to ensure the study's comprehensiveness.

Documentation of search			
Databased used	**Search terms**	**Total # of articles**	**Reviewed articles**
Google Scholar	"EU AI Act" and "governance" and "healthcare"	715	(Cancela-Outeda, 2024) – Cited by 97 (Butt, 2024) – Cited by 42 (van Kolfschooten & van Oirschot, 2024) – Cited by 125 (Murphy et al., 2021) – Cited by 469 (Chakraborty, & Karhade, 2024)
Google Scholar	"EU AI Act" and "governance" and "healthcare" and "low-income countries" or "developing countries" or "global south"	21 - 81	(Islam, & Wasi, 2025) – N.D. (Comunale & Manera, 2024) – Cited by 88 (Stankovich, 2025)

Carleton Omni Library	"EU AI Act" and "governance" and "healthcare" and "low-income countries" or “developing countries”	10	(Borines, Teckle, & Turi, 2025) (Xiong et al., 2025)
Reference list citation	N.A.	N.D.	(Aliman & Kester, 2019) – In (Birkstedt, 2023) (Al Zadjali, 2020) – In (Birkstedt, 2023) (Butcher & Beridze, 2019) – In (Birkstedt, 2023) – Cited by 278 (Cath, 2018) – In (Birkstedt, 2023) (Perry & Uuk, 2019) – In (Birkstedt, 2023) (Birkstedt, Minkkinen Tandon & Mäntymäki, 2023) – In (Comunale & Manera, 2024) (Lévesque, 2024) – In (Cancela-Outeda, 2024)

Limitations

Major limitations of this research are linked to the fact that the studied EU AI Act was adopted very recently (2024). It restricted the analysis to the content of the Act and existing critiques, rather than to evidence of how it has been applied in practice. The empirical evidence on how well the provisions in the Act are, are therefore non-existent. This means that my conclusions rest heavily on theory rather than on demonstrated results. Moreover, my research relying heavily on academic

critiques, commentaries, and interpretations give more room for researcher's biases. There is also very limited literature evaluating how specific healthcare needs are addressed as the Act is a more general framework.

Another major limitation is the treatment of LICs and LMICs as a single, relatively homogeneous group. This approach was necessary due to resource constraints and the chosen scope of the study, but it ignores important differences between countries that could significantly affect how the Act might be adapted or applied. As a result, some arguments may be less relevant to specific contexts, and the broad scope of the paper makes the findings more general and less tailored to particular situations.

Evolution of regulatory approaches

Countries are at different stages of developing AI governance frameworks, and they also adopt very different approaches that impact AI regulation in the healthcare sector. The most active nations in AI regulation are the United States the EU and China. Some Middle East regions are also picking up legislation and policy regulating AI in healthcare. One example would be India with its Indian Council of Medical Research (ICMR) which calls out ethical principles in that sector. Saudi Arabia has also its own Saudi Food and Drug Authority (SFDA) that elaborates expectations and requirements on AI/ML based device manufacturers, such as clinical evaluation, risk management and quality management systems. African nations are also picking up with a bigger focus on infrastructure development and user privacy. For instance, Rwanda has signed agreements with digital health companies for AI-based triage, symptom assessment, and cancer detection tools. (Chakraborty & Karhade, 2024)

Approaches also vary across regions. The EU and Brazil adopted a risk-based approach, the US adopted a more decentralized approach based on guidelines (FDA Regulatory Framework for Medical AI/ML), the UK adopted a context-based view with the implementation of Medicines and Healthcare products Regulatory Agency (MHRA) and the creation of the National Institute for Health and Care Excellence (NICE), China adopted algorithm recommendation and ethical reviews, while Japan and India have a more deregulated and flexible view on AI. (Chakraborty & Karhade, 2024; Comunale & Manera, 2024)

Moreover, the EU have always been a major actor in AI governance. The General Data Protection Regulation (GDPR) that was implemented in 2016 is recognized as the strongest privacy and security law in the world. It imposes obligations onto organizations anywhere in the world, only if they target or collect data related to people in the EU. (GDPR.EU, n.d.) It has influenced other regulatory frameworks such as Kenya's governance approach and South Africa's Protection of Personal Information Act (POPIA). (Islam & Wasi, 2025; POPIA, n.d.) In 2017, the EU implemented the Medical Device Regulation (MDR), which has also been very influential in shaping global standards for the safety, efficacy, and regulatory oversight of medical devices. (van Kolfschooten & van Oirschot, 2024) Other important examples of guidelines and codes that highlight principles for developing ethical AI are the European Parliament's resolution on Civil Law Rules on Robotics (February 2017), the European Union's Ethics Guidelines for Trustworthy AI (2019) and the European Commission's Proposal for a Regulation on a European Approach for Artificial Intelligence (2021). (Stankovich, 2022)

The illustration below shows that there is still very little AI legislation that is healthcare-specific, with only one framework that is both binding and tailored to the health sector. (Chakraborty & Karhade, 2024)

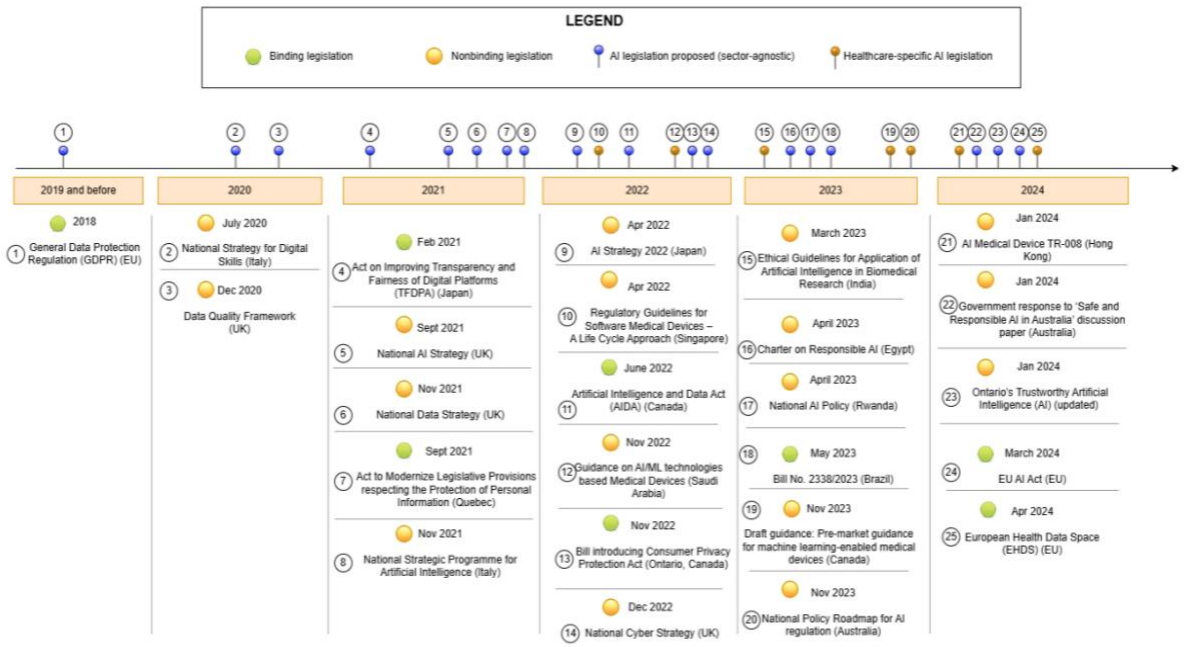

Figure 6: Timeline presenting the evolution of AI legislation across 14 jurisdictions, categorizing each legislation as binding or non-binding and sector-agnostic or healthcare-specific

Even though AI governance frameworks seem to be developing quickly across nations and regions all around the world. Most of the AI governance laws are often overarching and covering multiple sectors including healthcare. This horizontal approach carries many risks in health. It can make the regulations too broad, with cross-sector rules that often overlook the clinical realities of healthcare, like patient safety requirements, the need for rigorous medical validation, and the complexity of assigning responsibility when AI systems influence diagnosis or treatment. Therefore, specific focus on AI regulation in healthcare remains a major challenge that need to be better address.

What does good governance of AI in health looks like?

Regulatory frameworks on AI that are sector-specific to health should acknowledge international standards that are established by experts in that specific domain. The main ones are the WHO

Ethics & Governance of AI for Health (2021) and the OECD AI principles (2019). The WHO is a globally recognized authority that provides guidance on best practices and ethical principles for AI applications in health. To help with AI governance the institution identified five key areas of regulatory considerations for AI in health: documentation and transparency, data quality, risk management, privacy and data protection, and intended use and analytical and clinical validation. The illustration below shows which laws integrate each of these principles. The EU AI Act does touch on four out of the five principles. It does not address privacy and data protection as it already has the GDPR to cover this principle more specifically. (Chakraborty & Karhade, 2024)

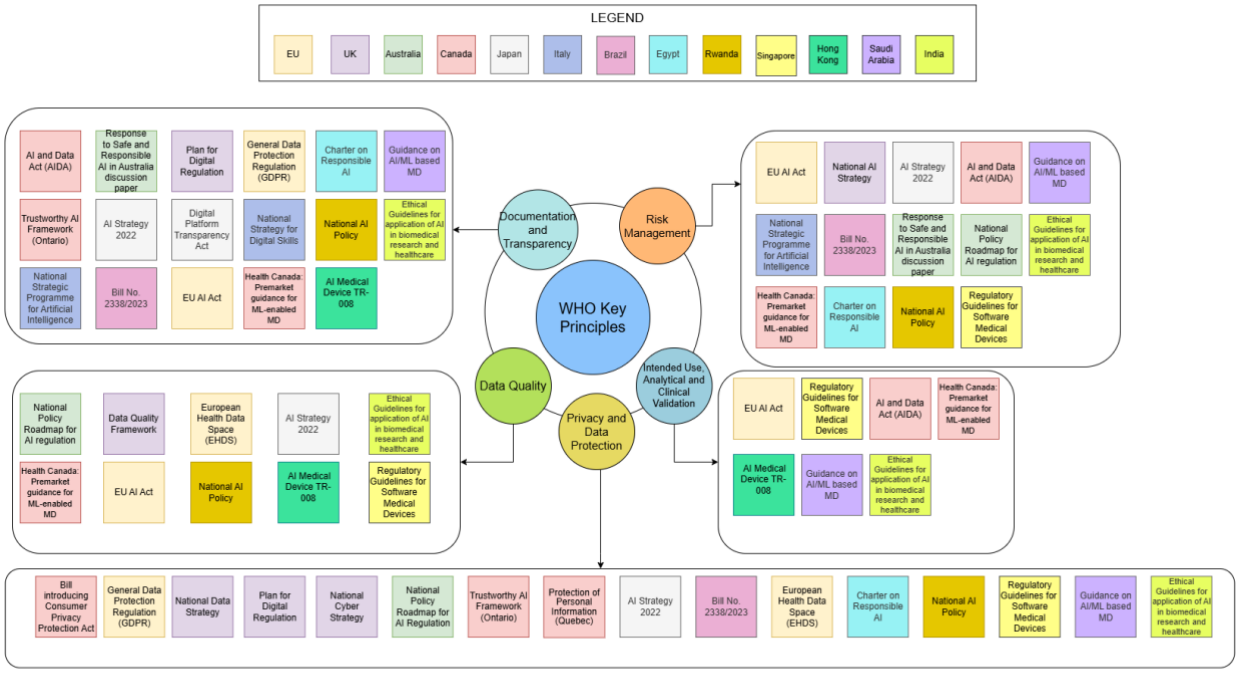

Figure 5: Depiction of laws illustrating WHO's core principles on ethical AI use in healthcare

The OECD AI principles were adopted in May 2019 and provide set standards for AI development and use that are practical and flexible enough to stay pertinent over time. The OECD principles include inclusive growth, sustainable development and well-being. They encourage AI to contribute positively to society and economic development. They also focus on human rights and

democratic values, making sure that AI respects fairness, privacy, and non-discrimination. Transparency and explainability are also prioritized so that AI decisions can be understood and challenged when necessary. Robustness, security and safety are also major elements that focus on minimizing risks and ensuring reliable performance in real world conditions. Finally, accountability ensures that developers, deployers, and users of AI remain responsible for the outcomes of AI systems, promoting trust and ethical oversight across all applications. (OECD, n.d.)

The European Union Artificial Intelligence Act (EU AI Act)

The EU Artificial Intelligence Act (AIA) was finalized in 2024 after many years of negotiation. It is the world's first comprehensive law that regulates AI and therefore aims to set global norms and put the EU as a leader in global AI governance. The Act was initially proposed by the European Commission in April 2021. It entered into force in August 2024 and will become fully applicable by August 2027. In the meantime, there are certain obligations, such as for high-risk AI that will come into effect as early as 2025. The act is part of the EU's broader digital policy framework and is aligned with the European Strategy on AI and the Digital Single Market Strategy. (Cancela-Outeda, 2024)

Objectives and scope

The main objectives of the AIA are to ensure that AI systems are safe and respect fundamental rights and EU values, to promote trust in AI technologies among citizens and institutions, to support innovation and competitiveness by harmonizing rules across Member States and to prevent harmful societal impacts and discriminatory or manipulative uses of AI. Since it is a regulation of

the EU, the AI Act applies directly and uniformly across all EU Member States. That way, it ensures that the implementation is consistent and homogeneous. The framework is based on a risk-based approach that categorizes AI systems into four levels. It is a horizontal structure, meaning that the risk categories are the same and apply the same across all sectors (healthcare, finance, education, transportation). The highest category is the "unacceptable risk" category which prohibits uses such as biometric surveillance or social scoring, meaning that these systems cannot be developed, deployed, or marketed in the EU. Below that, the "high risk" category includes AI systems used in medical devices or critical infrastructure, and requires strict obligations such as risk assessments, high-quality datasets, transparency, human oversight, and post-market monitoring, and these systems are subject only to light transparency requirements, so users know they are interacting with AI. The "limited risk" category covers applications like chatbots that must disclose their AI nature or wellness apps. Finally, the "minimal risk" category includes AI used for administrative tasks or research purposes, which face no specific obligations under the AIA and can operate freely. (Cancela-Outeda, 2024)

Collaborative Governance

One of the major innovations of the AI Act is its collaborative governance approach. This approach is "characterized by an ongoing exchange of knowledge between public institutions or agencies and diverse stakeholders (citizens, businesses, NGOs, experts, etc.) in policymaking" (Cancela-Outeda, 2024). It is an interesting approach as it promotes dialogue between diverse viewpoints that can create more informed policies, building on trust and legitimacy. It also helps create an environment that is based on collaboration and inclusivity in a world that needs it more than ever. Indeed, the complexity of AI and its fast development pace requires more collaboration and input

from multiple stakeholders to ensure that it is equitable, risk-free, and more. (Cancela-Outeda, 2024)

The governance system that is established under Chapter VII of the Act creates three main mechanisms that should facilitate collaboration among institutions, society, and other stakeholders. The goal is to use them as central facilitators to help with effective and consistent implementation. The first body is the European AI Office. It is the first entity ever to enforce binding rules on AI. It therefore aims to become an international reference in AI governance implementation mechanisms. It is created within the European Commission to supervise the enforcement and implementation of the AIA with member states. Its main tasks are monitoring compliance, developing risk assessment tools, coordinating with national authorities, and engaging in international cooperation on AI governance. National competent authorities remain the main institutions to oversee the implementation of AIA rules on their respective territories, but the AI Office ensure the coordination of the Act at the European level.

The second body that is created to generate collaborative governance is the European Artificial Intelligence Board. The board's role is to "advise and assist the Commission and the Member States in order to facilitate the consistent and effective application of this Regulation" (Cancela-Outeda, 2024). In other words, it serves as an advisory forum to promote coordination, share best practices, and support consistent application of the Regulation. It is connected to the European AI Office and is composed of representatives from each Member State.

The last body is the Scientific Panel of Independent Experts. This panel is composed of specialists in AI and related fields and provides technical and ethical expertise to guide enforcement and address emerging risks in AI. (Cancela-Outeda, 2024)

Limitations of the Act

The AI Act definitely marks a turning point in global AI regulation that will influence future regulations. It also introduces welcome safeguards for AI in healthcare, like stronger data governance and transparency requirements for high-risk systems. However, many critiques have emerged regarding the way it is presented. For instance, scholar Kolfschooten (2024) identifies four key limitations of the Act for the healthcare sector. Firstly, the lack of oversight and accountability for low-risk health AI could result in untested, harmful or ineffective systems, which could in return undermine public trust in AI. Secondly, the fundamental rights impact assessments are too ambiguous. The absence of a mandatory fundamental rights assessments for private healthcare providers could create disparities in patients' rights protection and the definition of "persons belonging to vulnerable groups" is too vague. Thirdly, the Act's exemptions for research and national security create loopholes for medical and biotechnological AI. Finally, he critiques the scientific research exemption as it could be too risky for patients, leaving room to escape rules for AI used in medical research and clinical trials.

Other scholars also argue that even though the collaborative governance framework aspires to inclusivity, it risks concentrating power in large industry actors and wealthier member states. (Lévesque, 2024) The uneven distribution of expertise and resources may limit equitable participation, particularly for smaller health institutions and marginalized groups.

Moreover, a horizontal approach has its limits in the healthcare sector as it applies the same categories and requirements across all industries. That approach is very different from the MDR sector-specific, technical framework tailored to healthcare. However, AI systems used as medical devices are expected to comply with both sets of rules. This mismatch is why experts often describe

the relationship between the EU AI Act and the EU MDR as an “arranged marriage” or even “conjoined twins”. The two frameworks were not designed together and therefore diverge when it comes to healthcare-specific regulatory needs. This example illustrates the limits of the EU AI Act’s horizontal approach in a highly specialized industry such as the medical device market. (Chakraborty & Karhade, 2024)

Xiong et al. (2025) findings also denounce the lack of structured frameworks in healthcare organisations that can translate the regulatory requirements into operational governance mechanisms from the EU AI Act. They explain that there is an urgent need for context-sensitive, operational AI governance in healthcare with a lack of practical implementation guidance for healthcare settings. Therefore, whether the Act can function as a true “gold standard” for AI in health will ultimately depend on how effectively it is implemented and adapted to real-world healthcare environments. When looking at LICs and LMICs, this becomes even more complex, as structural constraints may make it difficult to apply EU-level requirements or to benefit from them in the same way as high-income countries.

LICs and LMICs Context-Specific AI Governance

AI Landscape in Developing Countries

AI adoption is growing across LICs and LMICs, and especially in the health, agriculture, education, and public administration sectors. Health systems in these countries are often characterized by severe shortages of medical workers, limited equipment and resources, and a high disease burdens. WHO data shows that over 40% of countries have fewer than 10 doctors per 10,000 people, and that AI tools, such as clinical decision support systems, triage algorithms, and digital epidemiology platform could help relieve pressure on these human-resource shortages.

Indeed, early pilot projects have already shown that AI can help community health workers reach remote areas, improve early detection of diseases, and optimize medical resources. (Stankovich, 2022)

Murphy et al. (2021) mention in their paper the lack of research on AI ethics in LICs and LMICs and public health settings. They urge the need for further investigation into the ethical implications of AI in these contexts to ensure the feasibility of its ethical development and implementation on a global scale. The 2022 United Nations Development Programme's (UNDP) proposition of sustainable governance solutions for developing country governments of AI and big data deployment in healthcare also acknowledge the need for more research in that sector. Further mentioning that it is difficult to take technology innovation from developed countries and replicate them to address the needs of the developing world.

The UNDP proposition also identify eight potential uses of AI in health care that could further help determine what should be focused on in developing countries' contexts. AI can be used for training, research, end of life of care, treatment, decision making, diagnosis, early detection and keeping well. They highlight where AI could make the biggest difference for health systems that face chronic resource and workforce shortages and where governance frameworks must ensure safety, equity, and accountability. (Stankovich, 2022)

UNDP proposition of good governance in LMICs and LICs

Earlier in this paper, we have discussed the key principles of good governance of AI in health according to recognized international institutions such as the WHO and OECD. We also examined the differences between higher-income countries and lower-incomes countries in AI adoption and governance which are significant. Now, to understand further if the EU AI Act can truly be used

as a gold standard of good governance of AI for lower- and middle-income countries in the healthcare sector, it is important to examine the specific needs and challenges these countries face in that specific sector.

This research will use the 2022 UNDP proposition's to further analyze the specific needs of LMICs/LICs in governing AI in health. We will discuss basic regulatory capacity, accessible data and data quality, privacy and security, biases and inequality, human oversight, and public trust and community engagement.

Basic regulatory capacity

Governments will often lack resources and expertise to regulate AI technologies in healthcare. Moreover, they sometime have regulations that are outdated and unclear that can interfere and limit AI regulations. For instance, laws in Brazil, China and India require physicians and highly trained health workers to make diagnosis and carry out specific tests. These regulations limit which health services or advice can be given outside of a health facility or without the presence of a highly trained medical worker (which can now happen with the development of AI technology). To counter this, UNDP suggests developing clear guidance on when and where regulation on AI technologies is needed and to have at least a minimum of set principles and standards for data governance in data collection and data management. These two elements of data infrastructure are lacking in developing countries and would help ensure quality, interoperability and data-sharing in a way that preserves privacy. (Stankovich, 2022)

Accessible data and data quality

Health data is essential in health as it represents the basis of research and discovery of new treatments. It composes 30% of global stored data and it enables healthcare providers and

policymakers to make informed decisions about allocating resources. The COVID-19 pandemic showed the huge gaps that we still have in that area. This is even more flagrant in developing countries as health data is often incomplete or of low quality. According to OECD, many countries are still unable to extract and utilise the information they need to deliver better public health outcomes. Another challenge regarding health data in developing countries is the fact that they are often not interoperable or portable among institutions, leading to less diverse data sets that might not represent the patient population at the national level.

Beyond these structural issues, many developing countries also struggle to link existing data sources or use them for secondary research because data are siloed within institutions. This makes it harder to build the high-quality datasets needed for AI systems. In some cases, health data are not digitized at all or are captured by private sector actors who do not share them. When datasets are incomplete, poorly standardized, or unrepresentative, AI tools risk producing inaccurate or inequitable results. Therefore, a governance framework for AI in health in this context must ensure basic standards for data collection, management, and sharing. Clear national rules on data infrastructure would help reduce uncertainty for developers and support the safe deployment of AI tools. Ultimately, improving data accessibility and quality is necessary for ensuring that AI systems work effectively in low-resource settings and do not reinforce existing health inequalities. (Stankovich, 2022)

Privacy and security

A governance framework must ensure that health data are protected through clear privacy rules, security safeguards, and ethical data use standards. This is especially important in LMICs, where weak data protection laws and limited cybersecurity capacity increase the risk of data breaches, misuse, or commercial exploitation. While privacy issues could be addressed in other sections, the

UNDP mentions the importance of having an adequate and clear data protection systems for preserving data privacy specifically for health data. (Stankovich, 2022)

Biases and inequalities

AI systems tend to carry different types of bias when they are trained on Global North population and data because they don't understand the differences in LMIC and LIC populations. In AI design and deployment practices, we can find biases in real world patterns of health inequality and discrimination like biased clinical decision making, unequal access and resource allocation. There are also biases in the data itself because of sampling biases or lack of representative dataset. Then, there are biases in AI design and deployment practices as well as in its application because of power imbalances in agenda setting for instance. All of this results in a greater risk of discrimination, misdiagnosis, and exclusion of vulnerable groups that we find in majority in LMICs and LICs. Algorithmic bias is what we call biases in AI applications when created outside the developing world. That type of bias is considered to have more pronounced effect if that type of AI applications would be introduced to developing countries settings. Indeed, disease profile in low-income countries differ from those in developed nations. For example, in Sub-Saharan Africa, women tend to be diagnosed with breast cancer at a younger age and with more advanced disease than women in high-income countries. However, many diagnostic AI tools are trained on mammogram data from Europe, where the average patient is older and cancers are detected earlier. When these systems are deployed in Sub-Saharan Africa without adaptation, they can miss cases, delay diagnosis, and ultimately reinforce existing health inequalities.

So diverse and accessible data sets are not enough, they need to be representative enough of diverse populations. A new framework must require that AI systems are trained, tested, and validated on data that genuinely represent the populations they will serve. This means using diverse, locally

relevant and high-quality datasets; ensuring that algorithms are evaluated for bias before deployment; and involving local health experts, communities, and regulators in the design and review of AI tools. Independent national authorities should screen and certify AI systems to ensure they work equally well for all groups and do not reinforce existing health inequalities. AI developers should also be more transparent. That way, it would prevent damage but also help repair or compensate for harm that would be done in the worst-case scenario. It would also help with public trust by making stakeholders more accountable and help protect human rights in general. (Stankovich, 2022)

Human oversight

AI tools in health care should not fully replace clinicians, especially in low-resource settings because there are higher risks of misdiagnosis or inappropriate recommendations. The UNDP proposition mentions that maintaining a "human in the loop" is essential to protect patients and uphold basic standards of care. This means that adequately trained health professionals must validate AI-supported decisions and that AI systems should be designed to assist clinical judgement, not override them. Establishing strong oversight mechanisms is particularly important in developing countries, where regulatory bodies often lack capacity and where the consequences of low-quality AI tools can be severe. (Stankovich, 2022)

Public trust and community engagement

There is often a lack of trust in governments and institutions in the developing world that makes it harder to implement new regulations and policies. A new regulatory framework on AI in health would need to provide community engagement with the inclusion of clinicians, patients, and local actors in the design and deployment of the framework. Community engagement would therefore help address real health needs on the ground that would contribute to build that confidence that is

needed to successfully implement a new regulation. For example, data trusts are data-sharing frameworks that help organizations share data safely and fairly. They set clear rules about who can use the data, for what purpose, and how people's personal and sensitive information must be protected. They help shape the rules according to their specific needs. The WHO identifies data trusts as key model for stewardship of health data as they help build public trust, facilitate responsible data sharing among stakeholders, help rebalance power and provide a more equitable benefit sharing. Building trust is essential as many stakeholders in developing countries are concerned about who controls health data. (Boyd & Tennison, 2021) Since health data are often owned by the government, communities fear that private companies could gain access to them, exploit them for profit, or even move them outside national borders. A governance framework that clearly protects data rights, prevents monopolistic use, and ensures transparent, accountable data management is therefore crucial to earning public trust and ensuring safe and equitable AI adoption in health. (Stankovich, 2022)

What can be taken from the EU AI Act by LICs and LMICs

AI governance in healthcare is developing unevenly across countries. High-income regions like the EU, the US, and China have advanced regulatory frameworks, while middle-income and low-income countries are way behind. The EU AI Act represents the first comprehensive, risk-based regulatory framework designed to ensure safety, transparency, and accountability across sectors, including healthcare. Its collaborative governance approach, advisory bodies, and risk categorization could be a reference for structuring AI regulation and establishing mechanisms for oversight, stakeholder engagement, and technical guidance.

As seen in the past with other important AI regulation frameworks such as the GDPR or the MDR, they can be very influential in shaping new regulations in other countries including LMICs and LICs. However, replicating the EU AI Act as a whole is unrealistic as it was design without considering the specific reality of low-income settings in the developing world. Many countries face structural challenges such as limited regulatory capacity, shortages of technical expertise, fragmented institutions, weak data protection, and underdeveloped digital infrastructure that make complex frameworks difficult to enforce. Moreover, horizontal, cross-sector regulations like the EU AI Act may overlook healthcare-specific realities toward vulnerable populations and imported AI tools also risk exacerbating biases when algorithms are trained on Global North data that do not reflect local disease profiles and patient demographics.

However, some elements of the EU AI Act can inform and will probably influence LICs and LMICs governance frameworks as they address many of the specific challenges identified by the UNDP relatively to AI governance in health in lower-income settings. Firstly, the risk-based categorization of AI systems can guide prioritization of regulatory oversight for high-risk health applications. Secondly, requirements such as transparency, documentation and keeping a human in control of high-risk AI are essential for LMICs and LICs. Thirdly, the EU's emphasis on collaborative governance and stakeholder engagement provides interesting suggestions on how to implement an inclusive decision-making model that can support trust and community engagement. Finally, interoperability standards and data governance principles in the EU framework aligns with the need of robust health data systems to help with safe and effective AI deployment in limited resources context.

Although the EU AI Act includes many useful principles that can be used as benchmarks in AI governance in health, several parts remain not easily applicable in low-resources contexts. Even

in Europe, scholars have pointed out weaknesses that become even more critical in contexts with limited capacity. For instance, low-risk health AI is not closely monitored, fundamental rights assessments are vague, and exemptions for research or national security could undermine trust in AI. (Kolfschooten, 2024) Others have questioned its horizontal approach as well.

In LICs and LMICs specifically, the extensive documentation requirements may be too costly and resource intensive. The Act is also very dependent on robust regulatory institutions that most of these countries do not have. Finally, the GDPR's data protection requirements are highly technical and unrealistic to implement where privacy laws are already weak. This Act was designed for environments with strong enforcement mechanisms and is based on European sociopolitical priorities, therefore, it is important to use it carefully in other contexts. For most LMICs and LICs, the Act could be useful if they do a selective adaptation. They could integrate risk classification, transparency, human oversight and accountability with regional cooperation, and local and participatory bottom-up approaches.

Bibliography

Aliman, N.-M., & Kester, L. (2019). Extending socio-technological reality for ethics in artificial intelligent systems. *Proceedings of the 2019 IEEE International Conference on Artificial Intelligence and Virtual Reality, IEEE*, 275-282. https://doi.org/10.1109/AIVR46125.2019.00064.

Al Zadjali, H. (2020). Building the right AI governance model in Oman. *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance*, 116-119. https://doi.org/10.1145/3428502.3428516.

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. *Internet Research, 33*(7), 133–167. https://doi.org/10.1108/INTR-01-2022-0042

Borines, A., Teckle, P., Turi, A. N. (2025). Exploring the current AI landscape in Global South economies: A systematic literature review and research agenda. In A. N. Turin, P. Teckle, (Eds.), *Tech transformation and AI readiness. Pioneering paths for the Global South* (pp. 1-30). Springer Cham. https://doi.org/10.1007/978-3-031-73639-1_1

Boyd, M., & Tennison, J. (2021). *Health data governance summit. Pre-read: Data governance maturity and best practices*. Open Data Institute. https://cdn.who.int/media/docs/default-source/world-health-data-platform/events/health-data-governance-summit/preread-3-who-data-governance-summit_data-governance-of-health-data.pdf?sfvrsn=ffbd4a33_10

Butcher, J., & Beridze, I. (2019). What is the state of artificial intelligence governance globally? *The RUSI Journal, 164*(5-6), 88-96. https://doi.org/10.1080/03071847.2019.1694260.

Butt, J. S. (2024). Analytical study of the world's first EU artificial intelligence (AI) act, 2024. *International Journal of Research Publication and Reviews, 5*(3), 7343-7364.

Cancela-Outeda, C. (2024). The EU's AI act: A framework for collaborative governance. *Internet of Things, 27*. https://doi.org/10.1016/j.iot.2024.101291

Cath, C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*. https://doi.org/10.1098/rsta.2018.0080.

Chakraborty, A., & Karhade, M. (2024). Global AI governance in healthcare: A cross-jurisdictional regulatory analysis. *ArXiv*. https://arxiv.org/pdf/2406.08695

Comunale, M., & Manera, A. (2024). *The economic impacts and the regulation of AI: A review of the academic literature and policy actions*. (Working Paper No. 24/65). International Monetary Fund. https://doi-org.proxy.library.carleton.ca/10.5089/9798400268588.001

GDPR.EU. (n.d.). *What is GDPR, the EU's new data protection law?* Retrieved November 20, 2025, from https://gdpr.eu/what-is-gdpr/

Islam, R. M., & Wasi, A. T. (2025). C.A.R.E. for AI governance: Addressing regional challenges in the Global South. *Proceedings of sociotechnical AI governance opportunities and challenges for HCI workshop at the ACM CHI conference on human factors in computing systems (STAIG'25)*. https://chi-staig.github.io/papers/submission27.pdf

Lévesque, M. (2024). Collaborative AI governance: Minding the gap between theory and practice. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.4736131

Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., Lui, V., & Gibson, J. (2021). Artificial intelligence for good health: A scoping review of the ethics literature. *BMC Medical Ethics, 22*(1).

OECD. (n.d.). *AI principles*. Retrieved November 20, 2025, from https://www.oecd.org/en/topics/ai-principles.html

OECD. (2024). Explanatory memorandum on the updated OECD definition of an AI system. *OECD Artificial Intelligence Papers* (No. 8). https://www.oecd.org/content/dam/oecd/en/publications/reports/2024/03/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_3c815e51/623da898-en.pdf

Perry, B., & Uuk, R. (2019). AI governance and the policymaking process: key considerations for reducing AI risk. *Big Data and Cognitive Computing, 3*(2). https://doi.org/10.3390/bdcc3020026.

POPIA. (n.d.). *Protection of personal information act (POPI Act)*. Retrieved November 20, 2025, from https://popia.co.za/

Stankovich, M. (2022). *AI and big data deployment in health care: Proposing robust and sustainable governance solutions for developing country governments* (Background Paper No. 1-2022). United Nations Development Programme. https://hdr.undp.org/system/files/documents/background-paper-document/2021-22hdrstankovich.pdf

Van Kolfschooten, H., & van Oirschot, J. (2024). The EU artificial intelligence act (2024): Implications for healthcare. *Health Policy, 149*. https://doi.org/10.1016/j.healthpol.2024.105152

Xiong, H., Ledwidge, M. T., Fadahunsi, K. P., Lee, H., Wu, J., Morrow, S., Nisar, Y. B., Mbakaya, B., O'Donoghue, J., & Gallagher, J. (2025). Global artificial intelligence (AI) governance, trust, and ethics for sustainable health (GATES): A protocol for methodological framework. *Gates Foundation*. https://doi.org/10.12688/verixiv.1380.1