A principled governance for emerging AI regimes: lessons from China, the European Union, and the United States

Ren Bin Lee Dixon[1] (ID)

Abstract

Artificial intelligence (AI) governance is anticipated to have a transformative impact on humanity which has prompted researchers to analyze its implementation and use to ensure that the technology advances ethically and is beneficial for society. Though countries have begun to develop governance initiatives to regulate AI, the number of emerging AI regimes with an established structure is still relatively low. Meanwhile, the technology is advancing rapidly and has already caused harm inequitably to underrepresented communities. Thus, there is an urgent need to establish robust governance to mitigate the issues and risks attendant when deploying AI. While numerous ethics, principles, and structures have been recommended, this article intends to address the policy lag by providing policymakers with a simple and compelling AI governance framework that situates AI principles as the guiding baseline for developing and evaluating policies. Rather than devising new policy recommendations, the most recent (at the time of writing) and comprehensive governance documents from China, the European Union, and the United States were systematically selected, and examined in a comparative analysis to study how the three regimes address AI principles. Based on the comparative analysis, the most comprehensive and effective recommendations were selected to produce seven broad policy recommendations. The governance framework and recommendations are intentionally broad so that they can be adapted to adequately address AI principles across diverse contexts, encouraging the implementation of AI principles, increasing the likelihood of beneficial AI, and reducing the risks and harms associated with the technology. Nevertheless, the recommendations provided should not be considered exhaustive as the technology has an immense reach and new AI governance initiatives are developing continuously in this growth period in AI governance. It is thus essential for policymakers to survey the most current and relevant governance landscape to identify the best practices that are suitable for their specific context and need.

Keywords Artificial Intelligence · AI Principles · Governance · Policies · China · United States · European Union

1 Introduction

In the past decade, artificial intelligence (AI) has become increasingly more ubiquitous and has been deployed for a wide range of applications in diverse sectors. Increasing AI research has led to rapid growth of the technology. According to the Stanford Institute for Human-Centered Artificial Intelligence (HAI) 2021 AI Index Report, the number of AI publications in the world doubled from 2010 to 2021, growing from 162,444 to 334,497 [1]. In tandem with its research and technological advancement, various AI applications have become more affordable and accessible, pushing the average global AI adoption rate to increase by 6 percent from 2020 to 56 percent in 2020. Meanwhile, governance regimes on AI are emerging much more slowly—the 2021 AI Index reported the number of laws containing "artificial intelligence" grew from one in 2016 to 18 in 2021 globally—which has created a policy lag [2]. As a result, numerous issues and harm have been caused by the technology and it is affecting underrepresented subpopulations inequitably. The potential risks associated with the possibility of rapid and exponential growth in AI have further compounded global concerns regarding the safety of the technology [3, 4]. As a general-purpose technology that is expected to bring transformative changes across societies, there is an urgent need to establish robust governance frameworks to mitigate issues associated with AI development

✉ Ren Bin Lee Dixon
Dixon483@umn.edu

1 Hubert H. Humphrey School of Public Affairs, Minneapolis, MN, USA

Table 1 Topics and keywords for AI principles

Topics	Number of mentions	Keywords
For human	414	For human, beneficial, well-being, dignity, freedom, *diversity*
Fairness	374	Fairness, justice, bias, discrimination, prejudice
Transparency	348	Transparency, explainable, predictable, intelligible, audit, trace
Privacy	334	Privacy, data protection, informed, control the data
Safety	334	Safety, validation, verification, test, controllability, human control
Accountability	329	Accountability, responsibility
Security	223	Security, cybersecurity, cyberattack
Share	150	Share, equal, equity, power, *distributive*
Collaboration	103	Collaboration, partnership, cooperation, dialogue
Sustainability	55	Sustainability, *environment, Sustainable Development Goals (SDGs)*
Long-term AI	31	AGI, superintelligence, higher level AI

and deployment. This article intends to accelerate the development of governance regimes by proposing a simple and compelling governance framework grounded in AI principles to empower policymakers from diverse contexts to build principled AI policies. This method encourages the implementation of AI principles to shape a future of beneficial AI while minimizing AI's risks. Using the AI principle framework, the article provides a set of policy recommendations derived from a comparative analysis of governance initiatives from China, the European Union, and the United States.

2 Grounding governance in AI principles

While enforceable AI governance mechanisms such as regulations and policies have slowly begun to emerge in the last couple of years, the academia and both the private and public sectors have generated a wealth of non-binding AI principles and ethical frameworks in the past five years [5–7]. These principles and ethical frameworks proffer an abundance of well-researched guidelines that can be used as a foundation when developing AI governance frameworks. The research team lead by van Berkel et al. [8] found recurring discussions of ethical principles in national governance documents, suggesting a movement toward AI governance frameworks that are built upon a foundation of AI principles. Indeed, Raji et al. [9] proposed situating AI principles as the standard for evaluating the development of the AI life cycle and in internal audits when formalized guidelines are not available. By codifying compliance with AI principles into a risk analysis framework, this proposed method essentially implements the principles into practice.

Several papers have analyzed and distilled numerous AI principles documents into key topics, with some variations in the chosen representative words (for example, non-maleficence, humanity, beneficial, and freedom represent the

principle that AI should have a positive impact on humanity). The literature considered for determining the scope of AI principles in this article included Zeng et al. [12], Fjeld et al. [6], Floridi and Cowls [10], Greene et al. [11], and Jobin et al. [7]. Eventually, the website Linking AI Principles (https://www.linking-ai-principles.org/), created by Zeng et al. based on their 2019 paper, was selected as the main source of reference as it was the most comprehensive and updated source. On the website, Zeng et al. collected and analyzed a corpus of AI principle documents from 2016 onwards (see Table 3 in Appendix). The researchers first manually selected the core terms, then using a natural language processing algorithm, they identified and distilled down the keywords for each core term. The result is an overview of the key topics and their related keywords as shown in Table 1 [12].

AI principle topics and keywords based on Linking Artificial Intelligence Principles by [13]. This article includes several additional keywords (italicized) to address more current issues in cross-cutting policy areas.

For the analysis, this research included certain keywords expanding the topics to include more current issues in cross-cutting policy areas. For example, diversity was added under the principle "for human" to highlight the diversity of cultures, knowledge, and the rich context of human interaction that should be preserved in our interactions with AI. In developing a technology that aims to simulate human cognition and interaction such as AI, it is important to be aware of how AI will affect our humanness. According to Dick, humanness is an evolving concept that is relatively defined by our interactions, and in this case in relation to AI [14].

Distributive was included under the principle "share" to highlight the need to ensure that benefits and risks from AI are equitably distributed across societies. With its ability to convert resources into power (such as data to knowledge) with minimal effort, AI can easily concentrate wealth to a few large entities—especially to actors who deploy and

operate the AI system [15]. Therefore, the need to ensure that such power is equitably distributed becomes increasingly even more important.

Environment was included under the principle "sustainability" to underline the importance of upholding environmental health and sustainability throughout the development and deployment of AI. AI is positioned as a general-purpose technology that will trigger a transformative impact similar as the industrial revolution [16]. The industrial revolution has indeed been immensely beneficial to humanity and has propelled us into an era of wealth and growth, but the detrimental impact of the industrial revolution on the environment is undeniable and substantial. It would be unwise to neglect the lessons we've learned from this previous event. The SDGs from the United Nations were also included under sustainability to capture the globally endorsed mission to achieve "peace and prosperity for people and the planet, now and into the future" [17]. Finally, two keywords from the original table were removed, because they were considered more of a policy area (education) or because its definition was already covered by other keywords (confidential).

Many of the AI principles are interconnected—they depend on other principles to succeed and enable other principles to be realized. For instance, transparency in a recidivism risk analysis AI system would enable it to be audited to ensure that its decision-making processes are fair. Hence, when developing AI policies and regulations these principles should form a holistic foundation and not be siloed as individual criteria to be met [18].

3 Methodology

After defining AI principles as the grounding baseline for the governance framework, China, the European Union, and the United States were chosen as case references for developing the policy recommendations. The selection of these three regimes were based on their positions as current leading AI powers in the world, surpassing their peers in a series of progress indicators including number of AI talents, amount of research conducted, number of AI companies, adoption rate of AI systems, amount of data, and computing power [19]. These indicators amplified their positions as AI leaders on the assumptions they would have a relatively higher level of AI innovation, a wider extent of deployment, and more extensive experience with the impacts of the technology. Following these assumptions, these three regimes would have more studies on the impact of AI and the governance initiatives to address it. Additionally, their position as leading global AI powers make them influential actors in the global arena of governance setting [20–22]. As a consequence, their approach to AI governance will have significant influence over the emergence of AI governance

globally comparable to the "Brussels Effect," whereby emerging regimes would voluntarily extend the existing initiatives to inform their own and as a response to their relationship dynamics with these global AI powers. Once these three regimes were identified as case references, a systematic search was conducted to select a governance document from each of these regimes.

For the sake of compatibility and conciseness, only the most recent and comprehensive governance initiative developed by the main governing bodies from each regime were selected for the comparative analysis. Initiatives that were developed by think tanks, academia, or the private sector for specific applications or sectors were excluded. For instance, regulations and policies on data protection and privacy were excluded from the analysis even though data plays a huge role in AI. The objective is to compare and select the most effective and comprehensive initiatives in addressing AI principles, and to use that as the basis for the policy recommendations proposed in this article.

Based on the selection criteria for the governance initiative from these three regimes, the following documents were selected for the comparative analysis in this article [23–25]:

1. *Ethical Norms for the New Generation Artificial Intelligence* ("Ethical Norms") published by the People's Republic of China Ministry of Science and Technology (MOST) in 2021 [24]. The six-page translation of the Ethical Norms was included in its entirety in the analysis.

2. *A Proposal for Regulations in Artificial Intelligence* ("AI Act") by the European Commission in 2021. The AI Act is noteworthy for being the world's first bid at a comprehensive AI regulatory framework and is over a hundred pages long including detailed paragraphs under each article articulating regulatory implications [26]. Due to its length, only the Explanatory Memorandum section from pages 1–16 of the AI Act was included in the analysis [23].

3. the *Guidance for Regulation of Artificial Intelligence Applications* ("Guidance") released by the Executive Office of the U.S. President through the Office of Management and Budget. The Guidance was sixteen pages long and was included in its entirety in the analysis [25].

Once the representative governance initiative was selected from each governance regime, they were then compared and analyzed with AI principles using a matrix table (see Research notes in Appendix). This was then followed by an analysis of the effectiveness of their regulatory strategies to identify strength and gaps among the examined frameworks. Finally, based on the matrix table and analysis, a set of broad and amendable policy recommendations were developed to meet the AI principles.

Table 2 Policy recommendations for addressing AI principles

Policy recommendations	AI principles	Key points
Implement a centralized AI governance framework	For human, safety, sustainability	1. Comprehensive—applies to all sectors and applications, reduces regulatory gaps 2. Risk-based—proportionate regulations 3. Human- and environment-centered—human-centered design and environmentally sustainable
Establish robust data protection regulations	Privacy, safety, transparency, fairness	1. Comprehensive—to support a centralized AI governance framework 2. Data quality—preserve high data quality to increase fairness and accurate outcomes 3. Digital commons—once conditions are met, data sharing between private and public sectors can be promoted
Employ transparency as a compliance mechanism	Transparency, safety, accountability	1. Testing—provide clear reports on benefits and risks, and how AII is enabling and inhibiting AI principles 2. Audit trails—enforce traceability, encourage retrospective analysis of failures 3. Right-to-know—inform users of interactions with AI where relevant, such as when used in emotion detection and manipulative systems
Require testing to enforce safety and compliance	Safety, transparency	1. Low- and high-risk AI—while high-risk AI has a greater potential for harmful outcomes, certain low-risk systems can also cause harm to humans, such as biased search engine algorithms 2. Regulatory sandboxes—to balance regulatory restrictions and encourage innovation 3. Validation and verification—employ external oversight for objectivity, and include SDGs, international human right laws, and AI principles as validation framework
Collaborate with global alliances and local stakeholders	Collaborate	1. Globally—to accelerate global solutions, increase compatibility across borders, and prevent AI race 2. Locally—increase diversity in inputs for more equitable regulations, and to promote public trust
Invest in AI research	Long-term AI, security	1. Long-term impact—study the long-term advancement of AI and its impact on humans and the environment 2. AI governance—conduct impact assessment to support an evolving governance structure for emerging technologies 3. Cybersecurity—understand and anticipate novel threats and malicious use of AI
Implement distributive and redistributive policies	Share	1. Market balance—strong antitrust laws, prevent data monopoly 2. Society—implement progressive tax and explore basic income and negative income tax to address growing economic disparity 3. Pre-distributive policy—investigate its feasibility in developing economies to reduce global economic disparity accelerated by countries' ability to deploy and innovate AI

4 Policy recommendations for addressing AI principles

After completing a comparative analysis of the representative governance initiatives selected from China, the European Union, and the United States, this research developed policy recommendations based upon the most effective and comprehensive proposals among the governance initiatives produced by the three regimes (see Table 2). These policy recommendations are intentionally broad because, as a technology that will affect nearly all sectors and all walks of life, there is a need to consider the "big picture of what this will mean for ethics, governance, and societal impact" of AI [27]. The recommendations are intended to accommodate factors like the evolving nature of AI, the diversity in culture, legal systems, governance structure, and stages of AI deployment. Thus, specific recommendations targeting, for instance, each stage of the AI life cycle or any specific sector and application will not be discussed as such recommendations may differ depending on the factors mentioned above.

4.1 Implement a centralized AI governance framework

Given the impact AI already has on society and the transformative changes that are expected to take place, a key condition for ensuring that all AI principles are effectively incorporated throughout the technology's development and regulation is to lay the foundation for a unified and comprehensive governance framework. Policymakers must implement an overarching, centralized governance framework that establishes key definitions, standards, and regulatory approaches, similar to the AI Act. A centralized governance framework is crucial to enable interconnected AI systems to operate smoothly, and to better facilitate AI innovation across platforms. Having an overarching governance framework that establishes a consistent standard across sectors, local borders, and agencies will be highly advantageous for accelerating and regulating the deployment and advancement of the technology.

The central governance framework should adopt a risk-based approach similar to the AI Act, to categorize AI systems accordingly and enforce proportionate measures for regulating various levels of AI risks. While the AI Act and the Ethical Norms emphasized on AI's impact on humans, the risk-based approach should in fact assess the different levels of harm and the scale of impact an application could cause against not just humans, but the environment as well. By establishing clear risk levels, AI can be regulated proportionately according to its risk without stifling innovation—lower risk AI can enjoy more innovation and application freedom, while higher risk AI will be more regulated to avoid producing harm. For AI applications with dual uses—such as facial recognition for identification and verification—the method for determining risk levels should be treated with great care to monitor the evolving and emerging application of the systems, to avoid over- or under-regulating.

A centralized governance framework should determine the assessment parameters necessary to categorize AI systems into the appropriate risk levels; these levels can then be used by other agencies and business entities for developing their own specific AI regulations. The centralized framework can be understood as a minimum requirement for regulating AI, and by providing a clearly defined risk-based model, it can foster greater regulatory consistency across sectors. Additionally, the uncertainties regarding longer-term risks and benefits must be addressed using a precautionary approach because AI has a tendency to amplify its impact in intensity and scale, which could cause great harm to humanity [28].

Finally, the governance framework must be human-centered, as well as environmentally sustainable. While the Ethical Norms proposed human-centered AI design and the AI Act was framed to regulate AI harm against humans, a centralized AI governance framework must explicitly express that both regulatory efforts and AI development must be human-centered, as well as environmentally sustainable. This implies that regulations must incorporate or be founded upon relevant human rights laws (such as international human rights laws), and SDGs [17, 29]. These initiatives can be used as compliance frameworks when testing AI systems prior to deployment and during operations. Furthermore, human-centered AI design must preserve the rich context of human interactions and AI actors must be cognizant of how AI is redefining humanness [14]. The design and development stages of AI must include cross-disciplinary expertise from social sciences and humanities, and policies should ensure this expertise are included in AI education and recruitment processes.

4.2 Establish robust data protection regulations

To supplement a centralized AI regulatory framework, policymakers must establish a comprehensive and unified data protection governance regime to uphold individual privacy and encourage safe and secure collection, storage, and use of data. Though not analyzed in the research, all three governance documents referenced existing governance initiatives related to data protection. The China Personal Information Protection Law (PIPL)[1] and the EU General Data Protection Regulation (GDPR), for instance, are comprehensive

---

[1] Even though China's PIPL has clear provisions regarding how individuals and organizations handle the means of data processing, it is unclear how the specific provisions for the Chinese state government will impact user data protection [30].

initiatives that empower users by giving them the right to consent and providing users transparency on how their data is collected, processed, and used. The United States, on the other hand, has data protection laws that are sector- and state-specific, such as the Health Insurance Portability and Accountability Act, California Consumer Privacy Act, and Children's Online Privacy Protection Act. The U.S. segmented approach, however, is disadvantageous for protecting data in an age percolated with AI and big data. As each law is limited to its own domain, gaps and inconsistencies can appear across sectors and applications [31]. Additionally, since upholding individual privacy is one of the main objectives of data protection, extra attention must be given to accommodate the evolving concept of privacy as it responds to emerging technologies, and the meaning of privacy in different contexts and cultures [32].

Big data as a key component in AI will be generated at great velocity and in great volume, from various sources and in various forms. Therefore, a consistent and comprehensive data governance standard will provide a safe and secure environment to support the life cycle of big data across sectors, borders, and applications, which will also contribute to the acceleration of AI deployment and innovation. The AI Act and the Ethical Norms placed an emphasis on data quality to ensure safety, privacy, and fairness, especially since data plays an important role in machine learning training datasets. The quality of machine learning training datasets determines the robustness of an AI system and its compliance with AI principles. To ensure data quality, AI developers and researchers must implement measures to test for biases and inaccuracy that can lead to discriminatory outcomes in AI applications.

The availability of a robust data protection regulatory framework could facilitate the creation of a "digital common," as proposed by UNESCO [33]. A digital common can provide a secure space for private and public sectors to share the data they have collected with stakeholders, both for research and to further advance AI innovation. A digital common would also support the Guidance suggestion to increase access to granular government data, instead of aggregate data, to support AI advancement.

4.3 Employ transparency as a compliance mechanism

The principle of transparency was widely proposed as a compliance enforcement tool across all three governance initiatives from China, the European Union, and the United States. Policymakers must adopt this approach and require transparency practices involving high-risk AI and AI that

has the potential to manipulate humans to uphold safety and compliance to AI principles:

- AI operators must be transparent in their testing processes and required to produce clear reports on test outcomes so that users and stakeholders can be informed of the benefits and risks that may occur from using the application.
- Audit trails must be included in AI designs to enable traceability and collect information for retrospective analysis when failures occur [34].
- Design explanatory and exploratory user interfaces (such as mortgage loan applications where users can use sliding bars for adjusting income, assets, and loan amount) where relevant, to allow users to modify their inputs in AI applications and understand how different variables contribute to different outcomes [34].
- People must be provided with easy-to-understand statements explaining outcomes generated by AI that have consequential impacts on them. When necessary, these statements can be used for effective redress.
- Users must be informed when interfacing with AI that can manipulate people, such as AI that interacts with humans, that can detect emotions or reveal social categories, or manipulated content such as deepfakes.
- Where relevant, AI actors must explicitly implement AI principles throughout the life cycle of AI and provide transparent reports on which principles are enabled and inhibited.

Transparency, when executed appropriately, can potentially mitigate the information asymmetry caused by rapidly advancing AI technologies. Consequently, transparency can also foster greater public confidence in AI (as proposed by the Guidance) that will, in turn, increase the widespread adoption of the technology.

4.4 Require testing to enforce safety and compliance

While the AI Act proposed pre-deployment testing for high-risk AI, this recommendation should be extended to all AI systems that carry risks. AI systems with low- to high-risk levels must be tested prior to deployment, continuously monitored, and periodically tested throughout operation to ensure that the systems consistently comply with AI principles. Both low- and high-risk AI systems have shown evidence of adverse impacts on people, from algorithm biases in search engines to discriminatory outcomes in recidivism assessment applications [35, 36]. To uphold the standards of

these tests, external oversight entities should be established to prevent any conflict of interest. Regulatory sandboxes, such as those proposed by the AI Act, can also be used to test out new applications in a controlled and time-limited environment, without compromising on human and environmental costs.

Validation testing prior to deployment should assess AI's compliance with regulatory checks that include elements from the SDGs and international human rights, in correspondence to the AI principles. Presently, there are various options and methods available to test for fairness in AI systems, including toolkits for detecting and mitigating algorithmic bias, fairness-enhancing interventions, and building a collection of test cases to identify bias incidents [34]. But tools for testing AI's impact on SDGs and international human rights laws specifically still need to be developed. Similarly, verification testing must also be conducted on AI systems prior to deployment and over periodic intervals during operations, to ensure that outcomes remain consistent and as expected. A consistent outcome in AI is important as it also indicates that the system maintains its value alignment, supporting the long-term AI principle that powerful AI should be aligned with human values. Especially with AI's ability to self-improve and enhance its algorithms, continuous monitoring will be crucial to maintain consistently safe outcomes and detect any value misalignment as early as possible.

Clear and transparent reports on the outcomes of these tests must be made available to AI actors, users, and stakeholders. These reports could help AI developers identify which SDGs or human rights the AI system enables or inhibits and make the necessary adjustments to enhance the system's compliance with AI principles and regulatory requirements. Furthermore, continuous interval testing could help reduce failures and safety issues from occurring.

4.5 Collaborate with global alliances and local stakeholders

Countries should participate in global AI alliances to support the common advancement of AI, establish global governance, and attenuate the risk of an AI race. The AI principle of collaboration was recommended across all three documents examined in this article, with the objectives of promoting regulatory influence and technology advancement. By bringing diverse expertise together, global alliances can help accelerate AI innovation to address global challenges such as climate change, affordable clean energy, and greater access to quality education. Participation in global standard setting for AI will be essential to facilitate

a global AI ecosystem that encourages compatibility across borders and deter the development of malicious AI. Global alliances must implement measures to discourage rhetoric or intentions of an AI race and foster a strong common goal for advancing AI to benefit all of humanity.

Collaboration with local stakeholders will help policymakers take into consideration the many possible impacts AI can have on people. Policymakers should engage local stakeholders when developing AI policies and enforcing regulatory requirements, such as reporting AI failures. The Ethical Norms proposed easily accessible platforms by different groups of people and needs to facilitate active feedback and immediate reports on security vulnerabilities, regulatory vacuums, and policy lags discovered during the use of AI. Advocating for self-initiated feedback responses can encourage greater monitoring and regulation to achieve a comprehensive and agile governance framework. Actions and redress pertaining to failures should be transparent to keep AI actors accountable and incentivize people to report on failures consistently. These measures will foster greater public trust, uphold fairness, and provide equitable enjoyment of the benefits produced by AI.

4.6 Invest in AI research

Governments must provide continuous funding for research in long-term AI impact on people and the environment, potential cyber threats, malicious use of AI, and impact assessment on AI governance.

The principle of long-term AI advancement and its impact on people and the environment were not addressed in any of the governance initiatives examined, perhaps because it involves many uncertainties and the urgency to address other pressing issues at hand. While the exact timeline for the long-term advancement of AI is less certain, AI is anticipated to have an increasingly profound impact on societies that could be both advantageous and adversarial [3, 37]. By investing in research to understand its future potential, policymakers can take advantage of the coming opportunities to maximize benefits, while reducing policy lags that could cause substantial and irreversible harm. Additionally, AI governance is an emerging field with numerous proposed governance models that must be continuously examined to ensure it is effectively upholding AI principles, not unnecessarily impeding innovation, and able to keep pace with the rapidly growing application of AI [2].

All three governance initiatives recommended and referred to existing cybersecurity governance, however, the unique characteristics of AI warrant investigation into novel security threats and malicious uses associated with AI [38].

Robust security measures must be determined and implemented to prevent or manage these security breaches. Security as an AI principle is paramount for upholding safety in the deployment of AI, especially in safety components such as those in autonomous vehicles, machinery, and medical devices. Furthermore, as AI is increasingly adopted in public administration, the prevalence of cybersecurity as a national security threat will intensify correspondingly.

4.7 Implement distributive and redistributive policies

To counter the concentration of wealth and power produced by AI and uphold the principle of share, policymakers must implement distributive and redistributive policies to rebalance the scale and ensure that the benefits and risks generated by AI are distributed equitably.

Stronger regulations must be enacted in antitrust laws to prevent mergers and acquisitions that prohibit a competitive market. The nature of emerging technology as an evolving market element, combined with the limitless potential in AI, will require specific scrutiny on mergers and acquisitions in tech-related (e.g., fintech, biotech, healthtech, social media, etc.) companies. Specific investigation should be carried out to identify opportunities in emerging technologies that could lead to anticompetitive markets. For instance, Facebook's acquisitions of Instagram and WhatsApp were permitted at the time as the two smaller platforms were determined not in direct competition as Facebook. Presently, Facebook is being sued by the U.S. government for having become a social media monopoly and is being forced to sell WhatsApp and Instagram [39]. This case suggests that antitrust investigators must look beyond revenue and operational growth, and scrutinize the acquisition of data and users involved, considering them as resource and profit. The accumulation of massive amounts of data and users—which are positioned as two key resources in AI advancement—could contribute to a power monopoly through data monopolization. Indeed, when Facebook proceeded to integrate all three platforms into a unified structure, the move should have prompted more scrutiny as a potential anticompetitive practice.

The growing wealth inequality within and across countries is projected to be exacerbated by the implications of AI on the labor market [4, 40]. To slow growing wealth inequality, governments must implement progressive taxation on tech behemoths, such as taxation on monopoly rents and negative externalities created by AI deployment [41]. In the long run, economists expect societies to become wealthier as envisioned by Bostrom et al.—a future where AI reaches

a level of productivity that requires minimal gross domestic product [37]. In this regard, various basic income and negative income tax models should be explored, and governments must determine which models will be best suited for their economies, societies, and cultures. For instance, Korinek and Stiglitz recommended pre-distributive policies in developing economies instead of redistribution, because the "capacity to tax" in those economies will be low [41]. Additionally, a basic income strategy should be progressive and adaptable to the growing wealth generated by emerging technologies.

5 Non-AI principles policy considerations

Given its potential of having a transformative impact on societies, there will be other outcomes and implications that could arise from the deployment and use of AI. One of the implications of advancing toward a future permeated with AI is the need to improve AI literacy in the public so that individuals can continue to embrace their civic responsibilities and preserve their self-agency in a democratic society. Moreover, AI is also expected to transform the labor market as it offers greater efficiency, accuracy, and automation. Both matters can lead to fundamental shifts in societies that if left without any intervention could lead to harm against humanity. However, since they are beyond the scope of this article, policy recommendations advocating for AI literacy and the transformation of the labor market will not be included.

6 Limitations

There are limitations to focusing only on China, the European Union, and the United States, in the comparative analysis while omitting other countries and regions that have also been actively investing, developing, and deploying AI. The implications of examining only these three regimes are that the policy recommendations derived from this analysis will be to a degree skewed toward their unique realities, governance, and cultural context. Despite their advances in AI, all three regimes are at varying stages of implementing an AI governance framework. A consequence of this is that though the governance documents selected for this article meet the uniform criteria, the nature of the documents are not consistent and are thus weighted unevenly in the analysis. For instance, while the AI Act may have extensive recommendations involving the principle of transparency, the Ethical Norms offers only two simply because the AI Act is a

complete proposal for regulating AI while the Ethical Norms functions more as a general guideline. Additionally, these documents are relatively recent and there has not been substantial data and evidence on their effectiveness and impact on societies and AI innovation. Furthermore, the scope of this article is limited to analyzing the governance documents from these three regimes. The power dynamic between these three regimes plays a key role in their AI strategies and have a substantial impact on their governance initiatives, but this subject matter requires a much more extensive research and analysis, which is beyond the scope of the article. The analysis of this article solely focused on the three governance initiatives that are the Ethical Norms, the AI Act, and the Guidance.

Concentrating the analysis on these three governance initiatives also has its own limitations. All three regimes have established numerous governance documents on AI that are application or sector specific, supplementing the three documents examined in this article. Notably, these include the GDPR that governs data protection and privacy that came into effect in 2018, the U.S. National AI Initiative that contains an extensive list of AI strategy documents from national and federal agencies, and China released a 3-year road map for governing internet algorithms in 2021. By limiting the comparative analysis to these three selected documents isolates the analysis and neglects the fact that they are supported by an expanding governance framework. The distinct governance structures of the three regimes, their cultural context, and power dynamic have also informed their approaches to governing AI but are beyond the scope of the article.

The re-emergence of AI in the past decade suggests that the national governance of the technology is still in its infancy. Van Berkel et al. identified only 25 countries with an existing national AI governance framework, which they defined as including national policies and strategies. There is also an imbalance in the regions that AI governance research is coming from, with most literature originating from the global north that are also analyzing AI initiatives produced largely in the same region [8]. This could have implications on the interpretation of and the weight given to the AI principles that are used as fundamental guidelines for policies and regulations. The meaning and significance behind each AI principle may not be universal and there may be nuances among different cultures. Subsequently, a similar perspective could also be applied to the weight and approaches implied in the policy recommendations developed in these documents.

Therefore, the policy considerations that were based on the analysis of primarily developed nations may not necessarily be adequate for developing economies. In regions that are plagued by inaccessible healthcare, vulnerable to climate change issues, and threatened by geopolitical conflicts, risk levels and tradeoffs may be assessed differently to prioritize better health access, AI solutions for climate change issues, or advancement in military AI. Developing economies may have the opportunity to envision a more disruptive strategy and sustainable outcome through the deployment of AI. For instance, AI has the ability to assist developing nations leapfrog their healthcare system through applications in telemedicine as a way to mitigate healthcare worker shortages and extend healthcare to rural populations [42]. Hence, the proposed AI principle framework and the policy recommendations made in this article are intentionally broad for policymakers to use as reference or as a starting point for developing AI regulations and policies. In fact, the multitude of contextual considerations (cultural, historical, political, economic, and more) were impossible to be universally reflected in the policy recommendations. When policymakers adopt the proposed AI principle framework and policy recommendations, their distinct realities should inform their strategies for adapting the AI principle framework and recommendations to best serve their needs.

While this article attempts to address the transformative potential that AI will bring and its associated risks, it will not be able to address the myriad of benefits and risks emerging from different sectors and different applications of AI. There also isn't necessarily a universal approach to mitigate a similar group of risks in different cultural and application contexts. For example, people from individualistic cultures place a greater value on their privacy and are more resistant to data collection, while people from collectivist cultures are more likely to disclose their personal information for the benefit of the community [32]. In certain applications, facial recognition as verification on a single-person device may be less harmful than the use of facial recognition identification in a public space for surveillance purposes.

Finally, given that AI is a general-purpose technology, its impact will have such a vast reach that this article is unable to provide a comprehensive account of its full implications and thus it does not presume such a position. This article simply intends to offer policy recommendations that can help shape the outcome of AI that is aligned with the direction of AI principles. To direct AI innovation where its benefits are maximized to help humanity thrive in the long-term future, and its risks are minimized.

7 Conclusion

AI has become ubiquitous in our societies and will continue to permeate extensively, becoming an integral part of humanity in the future. Its impact on humanity is anticipated to be profound. It is crucial to set forth regulatory frameworks that can uphold AI principles continuously, to encourage a sustainable and beneficial development of AI innovations. Effective regulatory frameworks should maximize the benefits and diminish the risks of AI, and at the same time ensure that both benefits and risks are distributed equitably.

AI principles were positioned as the benchmark for the governance framework in this article, as they have been widely discussed, considered, and proposed in governance initiatives. The principles highlighted here were methodically condensed from a vast corpus of initiatives developed by the public, private, and academic sectors from around the world, representing a relatively wide and comprehensive contribution. The eleven key principles identified by the website Linking Artificial Intelligence Principles were for human, fairness, transparency, privacy, safety, accountability, security, share, collaboration, sustainability, and long-term AI [12]. By looking at the most recent and comprehensive governance initiative documents from current AI leaders—China, the European Union, and the United States—provided a reference point for developing policy recommendations by studying how these three regimes are addressing the AI principles. Based on the assessment, seven broad policy recommendations were produced to address the AI principles. The proposed AI principle framework for developing policy recommendations along with the set of recommendations were built to address the policy lag that is a result of rapid AI advancement, as well as to urge policymakers to implement AI principles when regulating and governing the technology.

Admittedly, AI governance is at present a developing domain and best practices from different realities and contexts are yet to be determined in certainty. Nevertheless, the governance initiatives analyzed in this article outlined several reasonable approaches based on existing AI governance regimes. For instance, China's requirements for human oversight throughout the AI life cycle, the European Union's risk-based approach, and the United States' wide-reaching recommendation for transparency in regulating and innovating AI. The comparative analysis indicated that a strong and comprehensive AI governance framework will be crucial for upholding AI principles and encouraging a safe and sustainable environment for advancing AI. Local and global collaboration in governance and innovation efforts will be immensely valuable for addressing the many risks and uncertainties that AI will bring. At the local level, AI stakeholders should have a role in shaping governance efforts that affect the outcomes of AI, empowering them to control the impact the technology will have upon them. Involving stakeholders in these processes reduces the risk of AI from overwriting their basic human rights. At a global scale, multilateral collaborations can help accelerate AI innovation to tackle global challenges, narrow the economic disparity among countries that could emerge from diverse AI deployment ability, and prevent the risk of an AI race.

Recommendations for addressing the principle of long-term AI were noticeably absent from all three initiatives, which perhaps reflects the current early stages of AI governance and innovation that prioritize addressing urgent issues such as safety, transparency, and security. Dafoe, however, urged for measures for governing long-term AI to be implemented today while "the stakes are relatively low" [43]. Hence, policymakers should seize the opportunity at present to shape the future into a space and time where we want to be; a space and time where humanity and the natural environment can thrive and progress sustainably. Equally important to explore today is whether we should embed human values into AI to ensure that our interactions with the technology and its impact on us will not detract our humanness as cautioned by Dick [14]. Hence, the decisions we make today will define not just the future outcome of AI, but also of humanity.

Appendix

See Tables 3, 4, 5 and 6.

Table 3 Overview of AI principles developed worldwide

Source	For human	Sustainabil-ity	Collaborat-ion	Share	Fairness	Transpare-ncy	Privacy	Security	Safety	Accountab-ility	Long-term AI
Beijing 2019	17	1	2	5	3	5	3	2	3	2	4
NGCNGAI 2019	10	1	4	5	6	3	6	1	7	4	1
AIIA 2019	11	1	1	3	7	7	4	6	10	4	
WHO 2021	4	3	1	8	12	23	12	2	11	16	
UNICEF 2020	5	1	2	2	4	6	2	1	4	2	
UNESCO 2021	48	8	3	9	16	29	17	5	9	12	
Telia 2019	2	1	3	2	6	4	1	1	4	5	
Smart Dubai 2019	13		3	1	4	4	6	6	7	2	6
OECD 2019	7	2	2	6	8	3	5	3	7	7	
NGCNGAI 2021	4	3	1	4	14	6	9	6	13	5	
Montreal 2018	6	3	2	7	3	4	8	2	5	7	
ITI 2017	10	1	5	3	3	1	3	8	6	12	
G20 2019	8	2	2	6	8	3	5	3	7	8	
FLI 2017	9		2	4	1	4	4	2	6	2	4
EGE 2018	15	2	3	5	9	2	10	5	8	6	
Beijing Children 2020	7	1	3	3	3	3	5	1	4	5	
Cabinet Office 2018	22	6	5	5	12	2	13	6	4	3	
EC 2019	2	2		4	11	9	6	3	9	5	
Sony 2018	3	2	1		2	2	2	2	1	1	
Australia 2019	9	1		3	5	4	6	5	7	14	
Shanghai YoungAI 2019	4	1		1	5	7	3	2	5	6	
SHAIISEAC 2019	2		1	1	2	2	4	11	7	3	
Russia 2021	12		3	2	7	3		2	5	9	
Nadella 2016	8		2	1	2	2	1	1	1	3	
MIC 2018	7		8	3	5	7	17	13	4	5	
MIC 2017	6		4	2	4	4	15	12	23	3	
Internet Society 2017	1		3	1	2	4	3	7	10	7	
UNI Global Union 2017	7	1		2	2	12	5	2	2	11	
HLEG 2018	15	3		6	35	13	11	4	10	8	
Google 2018	4		1	1	8	1	5	2	7	1	
US OSTP 2020	1			2	11	10	4	8	7	4	
Telefonica 2018	1	1	1		8	4	10	5	2		
Aotearoa 2020	4	1		1	5	5	3	2		2	
Deutsche Telekom 2018	4		4	2	4	4	4	6		9	
Tencent 2018	5			2	6	8	4	4	7		2
PDPC Compilation 2020	6	1		1	5	4		2	2	8	
Tsinghua CISS 2019	4		1	3	1	3	1	2	1		

Table 3 (continued)

Source	For human	Sustainabil-ity	Collaborat-ion	Share	Fairness	Transpare-ncy	Privacy	Security	Safety	Accountab-ility	Long-term AI
Montreal 2017	5		1	2	4	2	9		1	6	
JSAI 2017	8			2	5	1	3	2	7	3	
Intel 2017			1	1	4	2	8	5	6	3	
HAIP 2018	15		2	4	5		5		7	3	2
US OSTP 2020	1			2	11	10	4	8	7	4	
US IC 2020	3		1	1	1	3	2	5		2	
ICDPPC 2018	12		1	2	9	10	16		2	3	
IEEE 2017	13		1	1		8	5	3	6	8	
IA Latam 2019	3		4		2		2	1	2	1	
DoDDIB 2019	1			1	1	3		1	2	2	
SAP 2018	2		6		6	2	7	2	6		
The Public Voice 2018	1				12	4	4	5	8	13	
Tieto 2018	2			2	3	2		1	1	1	
GER DEC 2019	4	3		1	3		3	3	1		
CIGI 2018				1	4	5	3	3	12	5	
Megvii 2019	1				3	1	2	4	2	3	
COMEST 2019	2	1			1	4		1	1	3	
Alan Turing Inst 2019		2		2	4	3		1	1	3	
House of Lords 2018	4			1	4	4	2	3		1	
DeepMind 2017	1		4	1	1	2				2	
ITechLaw 2019	1				5	4	3		2	6	
Microsoft 2018					1	1	2	2	2	2	
Samsung 2019				1	3	3	1	1		2	
US DoD 2020				1	1	3		1	2	2	
Vodafone 2019	2				1	1	2	2		2	
Rome Call 2020	1				2	3	2	2		2	
NATO 2021	1				2	2		1	4	5	
PAI 2016	1		2			1	1	2		2	
OpenAI 2018	4		2	1				1	8		12
The Future Society 2017	4		2	2		3				3	
ADP 2018						4	8	1	1	1	
USACM 2017					4	4	2		5	3	
Unity 2018	2				2	1	1			2	
IEEE 2019	4					1	1		2	1	
FATML 2016				1	3	3	3			7	
Canada 2019				2		1	1	1		2	

Table 3 (continued)

Source	For human	Sustainabil-ity	Collaborat-ion	Share	Fairness	Transpare-ncy	Privacy	Security	Safety	Accountab-ility	Long-term AI
IBE 2018	1				5	5	5			7	
NYTimes 2019	1				2	2	1			2	
South Korea 2020	2						1	1	1	1	
TBS Canada 2018	1					2	1	1		1	
PDPC 2019	1				2	4			1		
Russia 2019	2		1			3		2			
Adobe 2021					4	2			1	3	
IBM 2018b	1				2	1				2	
IBM 2018a			1	1	1	5					
IBM 2017			1			1		1	2		
GE Healthcare 2018					1	2	1		2		
OP Financial 2018						2	4			1	
Baidu 2018	2			1					2		
US AI Initiative 2019							1	1	2		
Sage 2017	1				1					3	
Etzioni 2017							1	1			
Stanford 2018	1										
	414	55	103	150	374	348	334	223	334	329	31

*Descriptions in the source column link to their original documents. From Linking Artificial Intelligence Principles [12]

Table 4 Overview of key governance frameworks from China

Title	Year	Governing body	Description
New-Generation AI Development Plan (AIDP)	2017	The State Council of People's Republic of China	First national level legislative effort that explicitly focused on AI development as a unified strategy
Cybersecurity Law of the People's Republic of China	2017	The State Council of People's Republic of China	To ensure cybersecurity; safeguard cyberspace sovereignty and national security, and social and public interests; protect the lawful rights and interests of citizens, legal persons, and other organizations; and promote the healthy development of the informatization of the economy and society
Three-Year Action Plan to Promote the Development of a New Generation of Artificial Intelligence Industry	2018–2020	Minister of Industry and Information Technology (MIIT)	China's '3-year Guidance for Internet Plus Artificial Intelligence Plan (2016–2018)' focuses on: enhancing AI hardware capacity, (ii) strong platform ecosystems, (iii) AI applications in important socioeconomic areas, and (iv) AI's impact on society
AI Standardization	2018	China Electronics Standardization Institute	Outlines the national AI standardization framework and plan for AI capability development
Governance principles for the new generation artificial intelligence—Developing responsible artificial intelligence	2019	Ministry of Science and Technology (MOST)	This initiative highlights the theme of developing responsible artificial intelligence, emphasizing the eight principles of harmony, friendliness, fairness, inclusiveness, respect for privacy, security and controllability, shared responsibility, open collaboration, and agile governance
Ethical Norms for the New Generation Artificial Intelligence	**2021**	**Ministry of Science and Technology (MOST)**	**Lays out ethical norms for the use of AI in China. The norms cover areas such as the use and protection of personal information, human control over and responsibility for AI, and the avoidance of AI-related monopolies. The document does not specify how these norms are to be enforced; nor does it mention any punishments for those who violate the norms**
Internet Information Service Algorithmic Recommendation Management Provisions (Opinion-seeking Draft)	2021	Cyberspace Administration of China	To standardize Internet information service algorithmic recommendation activities
China Personal Information Protection Law	2021	Standing Committee of the National People's Congress	To protect personal information rights and interests, standardize personal information handling activities, and promote the rational use of personal information
Guiding Opinions on Strengthening Overall Governance of Internet Information Service Algorithms	2021	Cyberspace Administration of China Central Propaganda Department Ministry of Education Ministry of Science and Technology Ministry of Industry and Information Technology Ministry of Public Security Ministry of Culture and Tourism State Administration of Market Regulation National Radio and Television Administration	Over the next 3 years, to gradually establish a comprehensive algorithm security governance structure with a robust governance mechanism, a refined supervisory system, and a standardized algorithm ecosystem

The bolded highlighted document was chosen for comparative analysis. Sourced from Bal and Gill [20], OECD n.d. [44], Roberts et al. [45] and Sheehan [46]

Table 5 Overview of key governance frameworks from the European Union

Title	Year	Governing body	Description
General Data Protection Regulation (GDPR)	2018	European Union (European Union)	The GDPR is a regulation in European Union law on data protection and privacy in the European Union
European Union Strategy for Artificial Intelligence	2018	European Commission (EC)	The AI strategy proposed measures to streamline research, as well as policy options for AI regulation, which fed into work on the AI package
Policy and Investment Recommendations of Trustworthy AI	2019	European Commission (EC)	Provides recommendations that can guide Trustworthy AI toward sustainability, growth and competitiveness, as well as inclusion—while empowering, benefiting and protecting human beings
Data Governance Act	2020	European Commission (EC)	The instrument aims to foster the availability of data for use by making certain public sector data re-usable, increasing trust in data intermediaries, by promoting data altruism and by setting in place a governance mechanism for certain aspects of standardization
Digital Services Act Package	2020	European Commission (EC)	The new Digital Services Act package aims to modernize the current legal framework for digital services by proposing (i) clear rules framing the responsibilities of digital services to address the risks faced by their users and to protect their rights, and (ii) ex ante rules covering large online platforms acting as gatekeepers, which now set the rules of the game for their users and their competitors
AI Legislative Package (AI Act)	**2021**	**European Commission (EC)**	**The "AI legislative package" comprises: (i) a Proposal for a Regulation on a European approach for Artificial Intelligence; (ii) an updated Coordinated Plan with Member States, and (iii) a Proposal for a Regulation on Machinery Products. The draft legislation follows a horizontal and risk-based regulatory approach that differentiates between uses of AI that generate (i) minimal risk; (ii) low risk; (iii) high risk; and (iv) unacceptable risk, for which the EC proposes a strict ban**

The bolded document was chosen for comparative analysis. Sourced from the European Commission [47] and the OECD AI Policy Observatory [48]

Table 6 Overview of key governance frameworks from the United States

Title	Year	Governing body	Description
National AI R&D Strategic Plan	2018	National Science and Technology Council (NSTC)	Identifies the critical areas of AI R&D that require Federal investments
Federal Data Strategy	2019	Federal Geospatial Data Committee President's Management Council General Services Administration National Center for Education Statistics Department of Education and Training Federal Statistical Research Data Center Program Management Office U.S. Census Bureau Department of Commerce Federal Committee on Statistical Methodology Interagency Council on Statistical Policy Department of Education Office of Management and Budget Interagency Council on Statistical Policy	The Federal Data Strategy (FDS) encompasses a 10-year vision for how the Federal government will accelerate the use of data to deliver on its mission, serve the public, and steward resources while protecting security, privacy, and confidentiality
A Plan for Federal Engagement in Developing Technical Standards and Related Tools	2019	National Institute of Standards and Technology	
American Artificial Intelligence Initiative: Year One Annual Report	2020	Office of Science and Technology	This document provides both a summary of progress and a continued long-term vision for the American AI Initiative
Guidance for Regulation of Artificial Intelligence Applications	**2020**	**Office of Management and Budget (OMB)** **Office of Science and Technology Policy** **Domestic Policy Council** **National Economic Council**	**A memorandum that provides guidance to all Federal agencies to inform the development of regulatory and nonregulatory approaches regarding technologies and industrial sectors that are empowered or enabled by artificial intelligence (AI) and consider ways to reduce barriers to the development and adoption of AI technologies**
National Security Commission on AI	2021	National Security Commission on AI	The NSCAI Final Report presents an integrated national strategy to reorganize the government, reorient the nation, and rally our closest allies and partners to defend and compete in the coming era of AI-accelerated competition and conflict

The bolded document was chosen for comparative analysis. Sourced from National Science and Technology Council [49], National Security Commission on Artificial Intelligence [50]; OECD AI Policy Observatory n.d. [51], Office of Science and Technology Policy [52] and Vought [25]

Research Notes

Summarized definition of AI principles.

Matrix table of AI Principles with the Ethical Norms (China), the AI Act (European Union), and the Guidance (United States).

AI Principles: A Comparative Analysis of the Ethical Norms (China), the AI Act (European Union), and the Guidance (United States).

References

1. Zhang, D., Maslej, N., Brynjolfsson, E., et al.: The AI Index 2022 Annual Report (2022).
2. Taeihagh, A.: Governance of artificial intelligence. Policy Soc. **40**, 137–157 (2021). https://doi.org/10.1080/14494035.2021.1928377
3. Bostrom, N.: Superintelligence: Paths, Dangers, Strategies. Oxford University Press, Oxford (2014)
4. Brynjolfsson, E., McAfee, A.: The Second Machine Age, 1st edn. Norton Paperback, New York (2016)
5. Cath, C.: Governing artificial intelligence: ethical, legal and technical opportunities and challenges. Philos. Trans. R. Soc. Math. Phys. Eng. Sci. **376**, 20180080 (2018). https://doi.org/10.1098/rsta.2018.0080
6. Fjeld, J., Achten, N., Hilligoss, H., et al.: Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI. SSRN Electron. J. (2020). https://doi.org/10.2139/ssrn.3518482
7. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. Nat. Mach. Intell. **1**, 389–399 (2019). https://doi.org/10.1038/s42256-019-0088-2
8. van Berkel, N., Papachristos, E., Giachanou, A., et al.: A systematic assessment of national artificial intelligence policies: perspectives from the nordics and beyond. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society. ACM, Tallinn Estonia, pp 1–12 (2020).
9. Raji, I.D., Smart, A., White, R.N., et al.: Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. 12 (2020).
10. Floridi, L., Cowls, J.: A unified framework of five principles for AI in society. Harv. Data Sci. Rev. (2019). https://doi.org/10.1162/99608f92.8cd550d1
11. Greene, D., Hoffmann, A.L., Stark, L.: Better, nicer, clearer, fairer: a critical assessment of the movement for ethical artificial intelligence and machine learning. In: Proceedings of 52nd Hawaii International Conference System Science 10 (2019).
12. Zeng, Y., Lu, E., Ruan, Z.: Linking Artificial Intelligence Principles (LAIP). In: Link. AI Princ. LAIP. https://www.linking-ai-principles.org/keywords (2022). Accessed 19 Feb 2022.
13. Zeng, Y., Lu, E., Huangfu, C.: Linking Artificial Intelligence Principles. ArXiv181204814 Cs (2018).
14. Dick, S.A.: Making up minds. In: Think. Mach. Hist. Present Future AI. https://thinking-machines.online/dick/ (2021). Accessed 4 Dec 2021.
15. Calo, R.: Artificial Intelligence Policy: A Primer and Roadmap. Social Science Research Network, Rochester (2017)
16. Lynch, S., Andrew, N.: Why ai is the new electricity. In: Stanf. Grad. Sch. Bus. https://www.gsb.stanford.edu/insights/andrew-ng-why-ai-new-electricity (2017). Accessed 11 Mar 2022.
17. United Nations General Assembly.: Transforming our world: the 2030 agenda for sustainable development. In: U. N. https://sdgs.un.org/goals (2015). Accessed 19 Feb 2022.
18. Gasser, U.: AI and the law: setting the stage. In: Berkman Klein Cent. Collect. https://medium.com/berkman-klein-center/ai-and-the-law-setting-the-stage-48516fda1b11 (2017). Accessed 25 Jan 2022.
19. Castro, D., McLaughlin, M.: Who is winning the AI race: China, The EU, or the United States? 2021 Update. Center for Data Innovation (2021).
20. Bal, R., Gill, I.S.: Policy Approaches to Artificial Intelligence Based Technologies in China, European Union and the United States. Social Science Research Network, Rochester (2020)
21. Daly, A., Hagendorff, T., Li, H., et al.: AI, Governance and Ethics: Global Perspectives. Social Science Research Network, Rochester (2020)
22. Schumer, C.E.: S.1260—117th Congress (2021–2022): United States Innovation and Competition Act of 2021 (2021).
23. European Commission.: Proposal For A Regulation Of The European Parliament And Of The Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts (2021).
24. The National New Generation Artificial Intelligence Governance Specialist Committee.: Ethical norms for new generation artificial intelligence released. In: Center for Emerging Network Technologies. https://cset.georgetown.edu/publication/ethical-norms-for-new-generation-artificial-intelligence-released/ (2021). Accessed 25 Jan 2022.
25. Vought, R.T.: Guidance for Regulation of Artificial Intelligence Applications. Office of Management and Budget, Washington (2020)
26. Circiumaru, A.: Three proposals to strengthen the EU Artificial Intelligence Act. In: Ada Lovelace Inst. https://www.adalovelaceinstitute.org/blog/three-proposals-strengthen-eu-artificial-intelligence-act/ (2021). Accessed 7 Mar 2022.
27. Allen, J.R., West, D.M.: How artificial intelligence is transforming the world. In: Brookings. https://www.brookings.edu/research/how-artificial-intelligence-is-transforming-the-world/ (2018). Accessed 14 Feb 2022.
28. Armstrong, S., Bostrom, N., Shulman, C.: Racing to the precipice: a model of artificial intelligence development. AI Soc. **31**, 201–206 (2016). https://doi.org/10.1007/s00146-015-0590-y
29. McGregor, L., Murray, D., Ng, V.: International human rights law as a framework for algorithmic accountability. Int. Comp. Law Q. **68**, 309–343 (2019). https://doi.org/10.1017/S0020589319000046
30. Lee, A., Shi, M., Chen, Q., et al.: Seven major changes in China's finalized personal information protection law. In: DigiChina. https://digichina.stanford.edu/work/seven-major-changes-in-chinas-finalized-personal-information-protection-law/ (2021). Accessed 13 Mar 2022.
31. McGeveran, W.: Privacy and Data Protection Law. Foundation Press, Mineola (2016)
32. Li, Y., Kobsa, A., Knijnenburg, B.P., Carolyn Nguyen, M.-H.: Cross-cultural privacy prediction. Proc. Priv. Enhanc. Technol. **2017**, 113–132 (2017). https://doi.org/10.1515/popets-2017-0019
33. UNESCO.: Report of the Social and Human Sciences Commission (SHS). UNESCO (2021).
34. Shneiderman, B.: Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. ACM Trans. Interact Intell. Syst. **10**, 261–2631 (2020). https://doi.org/10.1145/3419764
35. Angwin, J., Larson, J., Mattu, S., Kirchner, L.: Machine bias. In: ProPublica. https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (2016). Accessed 31 Mar 2021.

36. Noble, S.U.: Algorithms of Oppression. NYU Press, New York (2018)

37. Bostrom, N., Dafoe, A., Flynn, C.: Public policy and superintelligent AI: a vector field approach. 28 (2018)

38. Brundage, M., Avin, S., Clark, J., et al.: The malicious use of artificial intelligence: forecasting, prevention, and mitigation. ArXiv180207228 Cs (2018)

39. Hamilton, I.A.: The FTC can move forward with its bid to make Meta sell Instagram and WhatsApp, judge rules. In: Bus. Insid. https://www.businessinsider.com/ruling-ftc-meta-facebook-lawsuit-instagram-whatsapp-can-proceed-2022-1 (2022). Accessed 15 Apr 2022.

40. Frey, C.B., Osborne, M.A.: The future of employment: how susceptible are jobs to computerisation? Technol. Forecast Soc. Change **114**, 254–280 (2017). https://doi.org/10.1016/j.techfore.2016.08.019

41. Korinek, A., Stiglitz, J.E.: Artificial Intelligence, Globalization, and Strategies for Economic Development. National Bureau of Economic Research, Cambridge (2021)

42. Yayboke, E., Carter, W.A.: The need for a leapfrog strategy. In: Cent. Strateg. Int. Stud. https://www.csis.org/analysis/need-leapfrog-strategy. (2020). Accessed 25 Feb 2022.

43. Dafoe, A.: AI governance: a research agenda. Gov AI Program Future Humanity Inst University Oxf Oxf 1442:1443 (2018)

44. OECD.: AI Policy Observatory Policy initiatives for China. In: OECD AI Policy Obs. https://oecd.ai/en/dashboards/policy-initiatives?conceptUris=http:%2F%2Fkim.oecd.org%2FTaxonomy%2FGeographicalAreas%23China. Accessed 28 Feb 2022

45. Roberts, H., Cowls, J., Morley, J., et al.: The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. AI Soc. **36**, 59–77 (2021). https://doi.org/10.1007/s00146-020-00992-2

46. Sheehan, M.: China's New AI Governance Initiatives Shouldn't Be Ignored. In: Carnegie Endow. Int. Peace. https://carnegieendowment.org/2022/01/04/china-s-new-ai-governance-initiatives-shouldn-t-be-ignored-pub-86127 (2022). Accessed 24 Jan 2022.

47. European Commission.: A European approach to artificial intelligence | Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence. Accessed 1 Mar 2022.

48. OECD.: AI Policy Observatory Policy initiatives of the EU. In: OECD AI Policy Obs. https://oecd.ai/en/dashboards/policy-initiatives?conceptUris=http:%2F%2Fkim.oecd.org%2FTaxonomy%2FOrganisations%23EuropeanUnion. Accessed 1 Mar 2022.

49. National Science and Technology Council.: The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update. National Science and Technology Council (2019).

50. National Security Commission on Artificial Intelligence.: The National Security on Artificial Intelligence. National Security Commission on Artificial Intelligence (2021)

51. OECD.: AI Policy Observatory Policy initiatives of United States. In: OECD AI Policy Obs. https://oecd.ai/en/dashboards/policy-initiatives?conceptUris=http:%2F%2Fkim.oecd.org%2FTaxonomy%2FGeographicalAreas%23UnitedStates. Accessed 1 Mar 2022.

52. Office of Science and Technology Policy.: American Artificial Intelligence Initiative: Year One Annual Report. The White House (2020).