Strengthening multi-stakeholder approach to global AI governance, protecting the environment and human rights in the era of generative AI

A report by the Policy Network on Artificial Intelligence

Disclaimer

The views and opinions expressed herein do not necessarily reflect those of the United Nations Secretariat. The designations and terminology employed may not conform to United Nations practice and do not imply the expression of any opinion whatsoever on the part of the Organization. Some illustrations or graphics appearing in this publication may have been adapted from content published by third parties. This may have been done to illustrate and communicate the authors' own interpretations of the key messages emerging from illustrations or graphics produced by third parties. In such cases, material in this publication does not imply the expression of any opinion whatsoever on the part of the United Nations concerning the source materials used as a basis for such graphics or illustrations. Mention of a commercial company or product in this document does not imply endorsement by the United Nations or the authors. The use of information from this document for publicity or advertising is not permitted. Trademark names and symbols are used in an editorial fashion with no intention of infringement of trademark or copyright laws. We regret any errors or omissions that may have been unwittingly made.

© Tables and Illustrations as specified

Suggested Citation Policy Network on Artificial Intelligence (2023). Strengthening multistakeholder approach to global AI governance, protecting the environment and human rights in the era of generative AI. Sipinen, M. (Ed.) United Nations Internet Governance Forum

This publication may be used in non-commercial purposes, provided acknowledgement of the source is made. The Internet Governance Forum Secretariat would appreciate receiving a copy of any publication that uses this publication as a source.

About the PNAI

This report is the first output document of the Policy Network on Artificial Intelligence (PNAI). The Policy Network addresses policy matters related to artificial intelligence and data governance. It is a global multistakeholder effort hosted by the United Nations' Internet Governance Forum, providing a platform for stakeholders and changemakers in the AI field to contribute their expertise, insights, and recommendations. PNAI's primary goal is to foster dialogue and contribute to the global AI policy discourse. Participation in and contribution are open to everyone. PNAI's recommendations and report will be presented and discussed at the 18th annual IGF meeting in Kyoto, Japan, in October 2023.

Authors

Chapter 1: Introduction and Chapter 2: Exploring AI and related aspects of data governance

Concettina Cassa (AgID, Italy); Rosanna Fanni (Centre for European Policy Studies, Belgium); Poncelet Ileleji (Jokkolabs Banjul, Gambia); Ricardo Israel Robles Pelayo (UNIR-EBC, México) Yashwinie Shivanand (Oxford Internet Institute, Canada)

Chapter 3: Interoperability of AI governance

Dr. Yik Chan Chin (Beijing Normal University, China); Sergio Mayo Macías (Aragon Institute of Technology, Spain) – Drafting team leads

Juan de Brigard (Fundación Karisma, Colombia); Celine Caira (OECD); Jin Cui (International Telecommunication Union (ITU, China); Sook-Jung Dofel (Deutsche Gesellschaft für Internationale Zusammenarbeit GmbH, Germany); Rosanna Fanni (Centre for European Policy Studies (CEPS), Belgium); Juan Manuel García (Derechos Digitales, Argentina); Zhengkun Hu (SenseTime, China); Jhalak M. Kakkar (Centre for Communication Governance, National Law University Delhi, India);Jochen Michels (Kaspersky, Germany); Umut Pajaro Velasquez (Independent researcher, Colombia); Tobore Oru (Digital Economy Forum DEF, Nigeria), Jamila Venturini (Derechos Digitales, Brazil); Prateek Sibal (UNESCO, India); Karine Perset (OECD); Gustavo Fonseca Ribeiro (UNESCO, Brazil); Carol Roach (Ministry of Economic Affairs, Bahamas); Yilin Sun (Peking University, China); Luis Aranda (OECD); Dr Allison Wylde (Glasgow Caledonian University, UK)

Chapter 4: Framing AI Lifecycle for race and gender inclusion

Umut Pajaro Velasquez (Independent Researcher, Colombia); Isabelle Lois (Federal Office of Communications OFCOM, Government of Switzerland, Switzerland) – Drafting team leads

Peter Martey Addo (Agence française de développement, Ghana); Tarcizio Silva (Mozilla Foundation, Brazil); Fernanda K. Martins (InternetLab, Brazil); Celine Caira (OECD, Canada); Gustavo Fonseca Ribeiro (UNESCO, Brazil) Helene Molinier (UNWomen, France); Allison Cohen (Mila - Institut québécois d'intelligence artificielle, Canada); Marielza Oliveira, (UNESCO, Brazil); Eleanor Sarpong (Impact Afrique Digital Partners, Ghana)

Chapter 5: Governing AI for a just twin transition

Shamira Ahmed (Data Economy Policy Hub, South Africa) – Drafting team lead

Mohd Asyraf Zulkifley (Universiti Kebangsaan Malaysia - The National University of Malaysia, Malaysia); Johannes Leon Kirnberger (OECD, Germany); Peter Martey Addo (Agence Française De Développement AFD, France); Florian Cortez, (Ethicqual, Netherlands); Muriel Alapini (Benin IGF, Benin)

Editor

Maikki Sipinen

List of contents

1. Introduction

In the 21st century, artificial intelligence (AI)[1] is considered a key driver of social and economic development. From smart homes and digital assistants to personalized learning or identifying medical conditions in CT-scans, AI applications are transforming every walk of life. If developed and deployed responsibly, AI can be used to for example to deliver more effective government services tailored to the needs of citizens, by improving transport services, health services, and infrastructure. Powered by quality data, AI contributes to cutting-edge innovations that aid technological development across sectors. Applications of AI and their impact transcend national boundaries and national or local interests, thus AI can be an essential tool in tackling pressing global challenges and accelerating the progress towards reaching the goals of the 2030 Agenda for Sustainable Development. AI can accelerate action aimed at improving social welfare, environmental stewardship, and sustainable economic growth.

While AI technologies can be of great service to humanity and all countries can benefit from rapid technological advances, the accelerated society-wide uptake of AI raises fundamental ethical concerns. For instance, biases embedded in AI systems can potentially result in AI systems that sustain and amplify existing unjust biases in our society, reinforce discrimination and enable new levels of authoritarian surveillance. Without decisive action and concerted interventions, AI could exacerbate discrimination, inequality, digital divides, exclusion and environmental harms, and deepen socioeconomic divides. AI can be leveraged to assist in the combat against climate change, for example in analyzing climate data, predicting climate patterns or optimizing energy use to identify vulnerable regions, assess risks, and develop strategies for climate adaptation. At the same time, AI's environmental impact derived from its intense energy consumption is a growing concern. Responsible AI and robust data governance can support climate adaptation and resilience efforts. Effective AI policy for the environment requires a fine balance between data governance that ensures high-value global datasets are accessible for responsible data usage to support public interest decision-making while reducing the environmental footprint of AI systems.

AI technologies have a great potential to be beneficial to the environment and society. However, for these benefits to be realized, the potential harms should not be ignored but addressed. It is vital to guide AI technologies, their development, uptake, and use, in a responsible direction. Many countries, regions and international organizations have developed AI strategies, policies, recommendations, regulations, and initiatives to maximize the benefits but also to manage the risks.[2] As AI's development and impact are global, international dialogue and joint action is needed. There is a need to continue developing,

---

1 There is no universally approved definition for AI, for the purposes of this report the writers draw from the AI definition developed by the OECD and used for example by UNCTAD and UNESCAP, that stresses the ability of machines and systems to acquire and apply knowledge to carry out intelligent behaviour. See for example: UNESCAP, *Artificial Intelligence in Asia and the Pacific* (Accessed 5.9.2023)

2 For national strategies see for example OECD.AI repository of AI policy initiatives: OECD.AI, *National AI policies & strategies repository* (Accessed 19.9.2023)

strengthening, sharing and implementing international recommendations, for example standard-setting developed through a comprehensive approach. These should place human dignity and human rights in the center and be grounded in gender equality, justice and mental well-being, diversity, interconnectedness, inclusiveness, and social and economic development - while also taking into account environmental and ecosystem protection.

The Policy Network on AI (PNAI) addresses policy matters related to AI and data governance.[3] It is a global multistakeholder effort hosted by the United Nations' (UN) Internet Governance Forum (IGF)[4], providing a platform for stakeholders and changemakers in the AI field to contribute their expertise, insights, and recommendations. The primary goal of the Policy Network is to foster dialogue and contribute to the global AI policy discourse.[5] This report is the first output document of the PNAI. It is not intended to be a comprehensive assessment or analysis of policy questions on AI and data governance. Rather, this first report develops analysis and recommendations to start a conversation. It delivers fresh suggestions from the global multistakeholder community, and paves way for the PNAI's future work on AI-enabled technologies.

Recognizing the opportunities and risks AI presents, the UN is promoting ethical development and application of AI and has for example committed to support AI-related capacity building for developing countries and broader stakeholder engagement on AI. The UN Secretary-General's Roadmap for Digital Cooperation presented in 2020 notes a gap in international coordination, collaboration and governance on AI, and calls for enhanced international multi-stakeholder efforts to ensure AI benefits all.[6] In July 2023, the UN Security Council discussed threats of AI to international peace and security for the first time. The Secretary-General announced the formation of a new high-level meeting on AI to assess options for global AI governance, as well as issuing new recommendations on AI governance to the UN Member States.

PNAI's work and this report contributes to the discussion on the topics of the UN's Global Digital Compact, a forthcoming agreement that focuses on the impact of digital technologies and their role in achieving the Sustainable Development Goals.[7] PNAI's recommendations and report will be presented and discussed at the 18th annual IGF meeting in Kyoto, Japan, in October 2023.

---

[3] IGF, *Policy Network on Artificial Intelligence* information webpage, Accessed 2.9.2023
[4] The IGF is a global multistakeholder platform that facilitates the discussion of public policy issues pertaining to Internet governance. For more information on see: IGF, IGF webpage, 2023
[5] PNAI, *PNAI Work Plan*, May 2023
[6] UN, *Roadmap for Digital Cooperation*, June 2020
[7] UN, *Our Common Agenda Policy Brief 5 A Global Digital Compact_an Open, Free and Secure_Digital Future for All*, May 2023

1.1. The IGF Policy Network on AI

PNAI work focuses on AI and related aspects of data governance. The policy network seeks to learn from and elevate AI governance frameworks, principles and policies being developed in and for the Majority World and Latinate languages, and to bring the IGF's multi-stakeholder community together, gather and synthesize knowledge in the community.[8] Participation in and contribution to PNAI are open to everyone. As a Policy Network under the IGF, PNAI seeks to build in-depth understanding of the topic, raise awareness and prompt cooperation across regions and stakeholder groups. The impact lies in the ability to facilitate discussion across stakeholder groups, facilitating a common understanding and inspiring and informing decision makers.[9] A Multi-stakeholder Working Group, consisting of experts, supports transforming the PNAI community's perspectives into actionable measures and recommendations.[10]

PNAI emerged from the request of the community: the Messages from the 2022 annual IGF meeting held in Addis Ababa conclude that the "IGF could be used as a platform for developing cooperation mechanisms on artificial intelligence. A policy network on artificial intelligence could be considered for the upcoming work streams in order to review the implementation of different principles with appropriate tools and metrics."[11] PNAI was launched in May 2023. The PNAI's work on AI and related aspects of data governance builds on previous discussions, reports, and the wealth of knowledge within the IGF community. Over the past years, the IGF has discussed topics including AI use by social media platforms and content moderation, dangers such as manipulation, deception, and mis- and disinformation, transparency needs in the operation and reporting of algorithmic systems, and necessary principles of rule of law, human rights, democratic values and diversity in the governance of AI.[12]

1.2. Multistakeholder process adopted to develop recommendations on AI

This report was developed through exploration and multi-stakeholder discussions in the PNAI community. The PNAI work and meetings are open for everyone to participate in. The information was shared through the PNAI website and open mailing lists. Invitations and updates were also shared widely through IGF mailing lists and social media channels as well as community updates through the IGF website. Further PNAI community members circulated invites and information of the work being done within their communities and stakeholder groups.

The work towards the report was structured in five phases: The first phase of the work was 'open dialogue', where the group defined three thematic focus areas for the report and agreed on a report outline. Early discussions on the PNAI held in spring 2023 highlighted the

---

[8] PNAI, *PNAI Work Plan*, May 2023
[9] IGF, *About the Internet Governance Forum*, 2023
[10] PNAI, *PNAI Multi-stakeholder Working Group* information website, 2023 (Accessed 8.8.2023)
[11] IGF, *Addis Ababa IGF Messages*, 2022
[12] For IGF publications and reports on past IGF activities, see IGF website

importance of focusing the dialogue and work on selected topics rather than striving to cover all areas relevant to AI. For the PNAI's first output report to bring value to the global AI dialogue, the aim should be to provide deep-dives areas that are central to fostering responsible AI development globally. Through an open brainstorming exercise, analysis and several commenting and input rounds the following thematic areas emerged: (i) Interoperability of global AI governance; (ii) AI gender and race; and (iii) AI and environment. The PNAI set up three sub-groups each dedicated to developing and drafting one of the three topical chapters of this report.

Once the topics were selected, the "information gathering" phase began. This took place through desktop research, engaging with invited expert speakers in the PNAI monthly calls[13] and by tapping in the expertise of PNAI members. 'Drafting the report' phase consisted of writing and editing the report together. The thematic drafting teams led the way and shared progress for comments and suggestions in the broader policy network meetings. The fourth phase of the process was 'consultation' where the draft PNAI report was shared with the wider IGF community for comments and suggestions. Finally, after editing based on the consultation input, the report was finalized and published to be discussed in the IGF 2023 annual meeting in October 2023.

---

[13] See summaries, presentations and meeting materials of all PNAI meetings on the PNAI website: PNAI, *Materials* information webpage, 2023 (Accessed 5.9.2023)

Figure 1: PNAI methodology
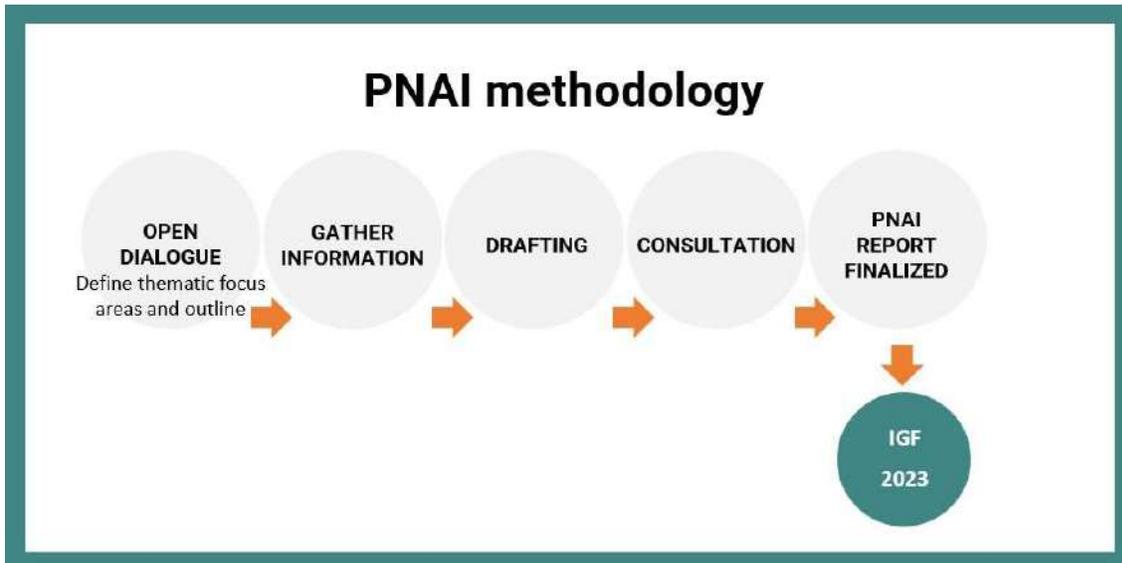
Figure 2: PNAI way of working
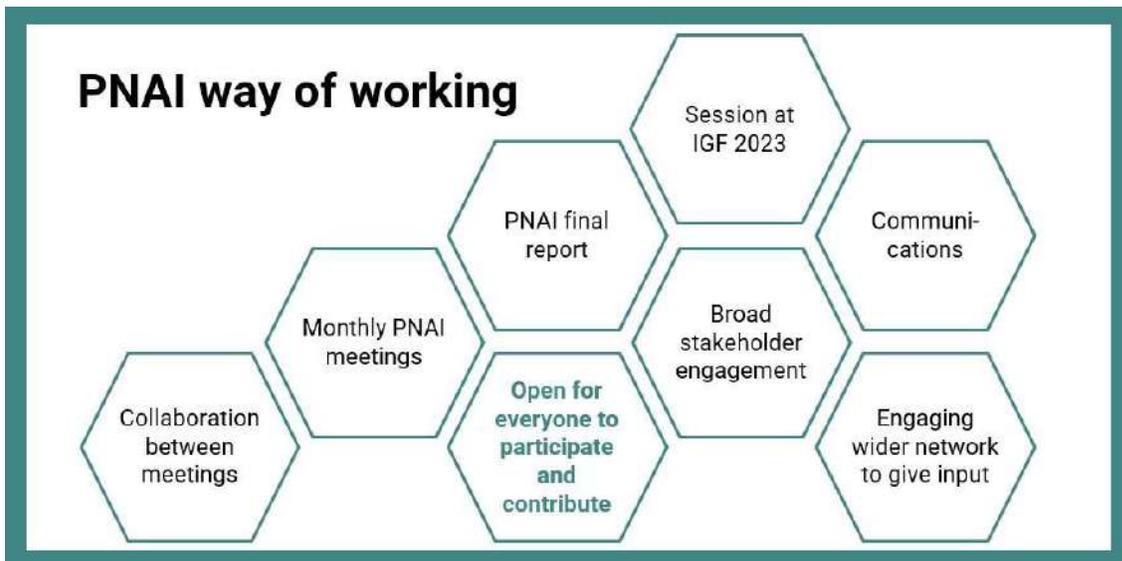
1.3. Structure of the report

The following chapter 'Exploring AI and related aspects of data governance' sets the scene and presents the Global South lens selected for the report. The chapter also describes the emergence of generative AI. The remainder of the report is structured around three thematic chapters each addressing one key topic relevant to AI and related data governance issues. Each chapter presents and assesses existing policy measures relevant to the topic, proposes next steps, and shares recommendations based on the multistakeholder discussions.

The first topical chapter, 'Interoperability of AI governance', delves to study the convergence and divergence among the different AI regulations being drafted by countries and regions; identify the AI development and policymaking gap and the challenges in strengthening global interoperability of AI governance towards AI that is secure, reliable, robust, fair, accountable and respects human rights and innovation. The chapter identifies and compares good practices and bottom-up initiatives that foster interoperability in AI governance and gives eight recommendations for further action.

The following chapter, 'Framing AI Lifecycle for gender and race inclusion', focuses on AI and gender, as well as AI and race. Do AI systems and harmful biases reinforce racism, sexism, homophobia and transphobia in society? Under which circumstances could AI be a force for good at improving gender and racial equality? What could be done to ensure that today's AI systems are a positive force in achieving that equality?

The third and final topical chapter 'Governing AI for a Just Twin Transition" takes a deep dive into the nexus of AI, data governance, and the environment, through the lens of two case studies. The purpose of this chapter is to comprehensively delve into the intricate interplay of AI, data governance, and the environment.

2. Exploring AI and related aspects of data governance

The PNAI community selected AI and related aspects of data governance as the topic of this report and the group's dialogues in the first months. The interdependence between AI and data is a critical nexus for addressing key societal challenges related to interoperability of AI governance, gender equity throughout the AI lifecycle, and environmental sustainability. Effective policies should strike a balance between fostering AI innovation and safeguarding the rights and well-being of individuals and the planet. By recognizing and acting upon this interdependence, we can harness the full potential of AI while ensuring a more equitable and sustainable future for all. Furthermore, robust data governance is needed to mitigate information asymmetries, ensure data quality, and address multidimensional power dynamics to drive informed climate adaptation, resource management, and conservation efforts, essential for addressing the biggest challenges of our time.

PNAI's work on this report started from the observation that there is a plethora of AI governance frameworks, ethical AI policy approaches, documents, strategies, and forums, but the vast majority of these have been developed in or for the Global North.[14] The policy network set as one of its goals to look at AI and related aspects of data governance from the Global South Perspective. Inspired by the leaps in technological development that have dominated the headlines at the time of writing this report, PNAI decided to dive deeper into the world of generative AI technologies. Setting the scene for the report and its recommendations, this chapter presents the Global South lens selected for the report and provides an introduction to the world of generative AI.

2.1. Viewing AI policy debates through the Global South lens

The Roadmap for Digital Cooperation issued by the United Nations Secretary-General in 2020 noted a lack of representation and inclusiveness in the international coordination and collaboration on AI.[15] The Roadmap's call for diverse stakeholder participation in global digital cooperation is particularly relevant if we consider the underrepresentation of Global South[16] countries in the drafting of AI principles. A 2019 study by ETH Zurich researchers[17] found that the USA, UK and Japan alone were responsible for most of the 84 ethics and AI documents identified for analysis. Although the sample analyzed then does not represent the current landscape of AI standards and guidelines, it is clear that certain countries and regions are

---

[14] IGF Multistakeholder Advisory Group, *Proposal for an IGF 2023-2024 Policy Network: Policy Network on Artificial Intelligence,* February 2023

[15] UN, *Road map for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation*, May 2020

[16] As synthesized Haugh: "As a meta category, the 'Global South' has taken on a variety of meanings. It refers not just to landmasses and waters south of the equator, the strictly defined hemispheric south. Instead, the term has been a general rubric for decolonised nations roughly south of the old colonial centres of power.". For more details on the main meanings connected to the concept of Global South, see: Sebastian Haug, *What or where is the 'Global South'? A social science perspective*, September 2021

[17] Anna Jobin, Marcello Ienca and Effy Vayena, *The global landscape of AI ethics guidelines*, Nature Machine Intelligence, September 2019

responsible for most of the global dialogue and development in this area. AI ethics principles do shape policy debates at global, regional and national levels, but oftentimes, such supposedly "global" processes ignore contextual particularities including concerns and needs of the Global South.

In addition to the underrepresentation of Global South in AI policy documents identified at global level, it is important to shed light to the questions of inclusiveness and representation when developing AI policies in and for regions or countries. Evidence from Latin America, for instance, shows that public participation was limited in the drafting processes of national AI strategies. The processes in general failed to involve the groups which can be most affected by such technologies. This is the finding of a 2022 study that describes the scant participation of women and the failure of government institutions to provide disaggregated data which would demonstrate the representation of priority groups.[18]

The findings of the studies are in stark contrast with the ongoing vibrant discussions on global AI governance, numerous initiatives to develop AI grounded in justice and equality and the research undertaken on the topic in Global South countries. It also doesn't take into account the key role such countries play across the AI value-chains, for example as providers of minerals that are fundamental for the development of their infrastructure, energy to sustain data centers, data and workforce[19] to train algorithms or as final users of systems. The exclusion of Global South countries from policy debates on AI invisibilizes key priorities from discussion.

Global South participation in global AI policy debates is key. As highlighted in several international standards, AI governance, development and deployment should be discussed in different organizations, groups, parts of the world by experts, enthusiasts and laymen with different backgrounds. According to the UNESCO's AI Ethics Recommendation, participation of different stakeholders throughout the AI system life cycle is necessary for inclusive approaches to AI governance, enabling the benefits to be shared by all, and to contribute to sustainable development.[20] Otherwise, global inequalities between North and South tend to increase, as the AI industry is concentrated in a few developed countries and their systems are built from the extraction of value from less developed regions, including Africa and Latin America. Thus, building frameworks which guarantee sustainable, human rights-compliant AI requires both North-South and South-South collaborations.

---

[18] See research report by Derechos Digitales: Laura Hernández, María Paz Canales and Michel de Souza, *Artificial Intelligence and Participation in Latin America: the national AI strategies,* 2022

[19] See for instance the precarious work conditions of people training AI systems in Global South countries: Niamh McIntyre, Rosie Bradbury and Billy Perrigo, *Behind TikTok's Boom: A Legion of Traumatized, $10-A-Day Content Moderators*, article in TIME magazine, October 2022; Veronica Smink*, Los cientos de miles de trabajadores en países pobres que hacen posible la existencia de inteligencia artificial como ChatGPT (y por qué generan controversia)*, BBC News Mundo article, March 2023

[20] See the ethical AI framework that has been adopted by 193 countries: UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2022

The PNAI community seeks to learn from and elevate AI governance frameworks, principles and policies being developed in and for the Global South and non-latinate languages. With our growing global network, we can bring value to the AI dialogue by leveraging Global South perspectives, which are vital but typically missing or under-represented on AI policy debates. Hatched under the IGF network, PNAI can build on two decades of experience organizing global, multistakeholder discussions on digital governance. It can also benefit from the IGF's concrete mechanisms for engaging the Global South through its more than 155 national and regional initiatives.[21]

At the same time, PNAI acknowledges the several imbalances that prevent Global South stakeholders from having a meaningful participation even in spaces built for inclusive worldwide participation, such as the IGF. These include limited funding to travel and precarious connectivity conditions to participate in events, the prioritization of English as the main language, among others. Global multilateral organizations committed to opening spaces for multistakeholder participation should take into account such inequalities in their design in order to foster true global dialogue and to ensure Global South perspective is included.

2.2. In the wake of generative AI

Generative artificial intelligence has emerged to form one of the most promising and, at the same time, most controversial areas in AI development. Until recently, machine learning was mostly limited to predictive models (analyzing data to make predictions) while generative AI is a specialized branch of AI that focuses on learning from various data patterns with the purpose of creating new content. Systems powered by generative AI, such as Open AI's ChatGPT and GPT-4, Anthropic's Claude, or Google's Bard, create texts, images, videos, music, software design, or scripting for test codes based on prompts by the user. Due to its versatility, generative AI is increasingly employed across different areas including economy, social interaction, business, arts, and academia. These tools can tackle repetitive tasks swiftly and efficiently, leading to a significant boost in productivity. Generative AI is expected to increase productivity across sectors, estimates show it could add USD 2.6 to 4.4 trillion annually to the global economy.[22]

Generative AI carries the potential to benefit or harm vulnerable groups and communities. On the one hand, it makes possible personalized solutions. Generative AI can for example help teachers create personalized lesson plans for each student.[23] It is already assisting blind or low-vision people by turning images into text interpretation in numerous languages.[24] These and many other linguistic or cultural adaptations make services more accessible. However, if the data used to train generative models is not representative, generative AI could perpetuate stereotypes and biases, exacerbate discrimination, and increase inequality. AI developers and

---

[21] IGF, *About the Internet Governance Forum*, 2023
[22] McKinsey, *Economic potential of generative AI*, June 2023
[23] Kevin Roose, *Don't Ban ChatGPT in Schools. Teach With It* article in New York times, January 2023
[24] OpenAI, *Be My Eyes* customer story information webpage, March 2023 (Accessed 12.9.2023)

society in general must make a conscious effort to ensure that generative AI is developed and used in ways that empower underprivileged groups, rather than further marginalize them.

In the case of generative AI and gender, there are high expectations to design algorithms that allow raising awareness on the topic. Therefore, the objective should be not only to use generative AI as a tool for the study, analysis, and promotion of gender issues but also to guarantee that these systems are trained with accurate and representative data linked to awareness, avoiding false content or information with discriminatory visions, thus contributing to a broader and fairer understanding of gender issues in today's society.

The ability of generative AI to generate content, such as text and images, raises serious ethical concerns. It can be used for disinformation and other forms of digital manipulation. Speaking to the Security Council in July 2023, the UN Secretary-General highlighted the capabilities of new generative AI models, and warned about the risks that the advent of generative AI can bring, for example for disinformation and hate speech.[25] Furthermore, the integrity of the information and the protection of personal data and individual privacy are at risk. It is essential to establish clear limits and regulations that protect individuals from possible abuse, without going against innovation.

Generative AI has the potential to democratize digital services, as it allows the creation and adaptation of content in an automated way.[26] However, if not managed properly, it could lead to digital monopolies where a few companies control access to and use of generative technology. A truly open and free digital future demands that generative AI be developed and distributed in a transparent, equitable and accessible manner.

The digital age has led to the emergence of new challenges, including manipulation, deception, and misinformation on the Internet. Although it can be a potential source of problems, generative AI could also provide solutions to combatting these same challenges. Systems building on this technology are used to create text, images, or videos, they can also be developed and trained to help detect false information, for example fake news or deepfake images, generated by the powerful technology. Researchers and developers are working to develop tools that would allow tagging fake content when it's being created or detecting fakes after they have been published.[27] Disinformation can spread and influence public opinion at an alarming speed, therefore such tools need to be urgently developed and implemented.

Another concern to consider is fragmentation on the Internet, where algorithms personalize and limit the information that users have access to. This distinction can be counterproductive, as it reinforces existing opinions and limits exposure to diverse perspectives. Generative AI can be trained and used, to analyze broader patterns and understand context. Moreover, it can

---

[25] UN, *Secretary-General Urges Security Council to Ensure Transparency, Accountability, Oversight, in First Debate on Artificial Intelligence,* press release, July 2023

[26] For examples of generative AI's impact on public services, see: CoE General Secretariat Analysis and Research team, *ChatGPT in the Public Sector − overhyped or overlooked?,* April 2023

[27] Nicola Jones, *How to stop AI deepfakes from sinking society — and science,* Nature news feature, September 2023

contribute to the creation of more balanced filters that present information in a more impartial way. Finally, human rights can be affected or violated by misinformation or the construction of misleading narratives. There is a risk of using this technology as a surveillance or repression tool by authoritarian regimes. Therefore, it is necessary to design an ethical framework and its implementation.

Generative AI has the potential to transform industries and society, to boost innovation across diverse fields, from arts to scientific research and empower individuals including marginalized groups. To ensure generative AI contributes towards a positive future, it is crucial to prioritize responsible design and release practices from the beginning. As generative AI continues to advance at an unprecedented pace, there is a need for collaboration among stakeholders to ensure that AI serves as a force for good. The IGF Policy Network on AI promotes the debate on how to increase international cooperation among the stakeholders on the use of generative AI and related aspects of data governance.

2.3. Global multistakeholder dialogue is crucial in getting global AI governance right

Understanding AI's future impact on societies is very difficult. Governing and regulating a technology in development is difficult but it is likely to become even more difficult later when the technology becomes deeply entrenched and its effect on society is better understood. Under these circumstances, making effective and informed decisions on AI is complex. Bringing in diverse perspectives and expertise can enhance understanding of the implications of AI in a holistic manner, and it is necessary for developing relevant and applicable policy for the national and international context. Multistakeholder approach facilitates the development of inclusive AI policies that help decision makers to consider diverse viewpoints and expertise, prevent capture by vested interests, and counteract polarization of policy discourse.

The multistakeholder dialogue is crucial for addressing AI's policy evolution. But it is not easy to create spaces for truly global AI dialogue or reach stakeholder groups, for example it can be challenging to include grass-root organizations with limited financial means in the global discussions. Already in 2020 there were over 160 organizational, national, and international sets of AI and governance principles worldwide[28] but so far, no common platform to bring these initiatives together. At the time of writing this report, PNAI is in its early stages, but is already bringing value to dialogues on AI governance as it is an open forum that brings diverse stakeholders from across the world together for timely discussions on AI. The impact of IGF's intersessional activities, such as PNAI, comes from facilitating global discussion across stakeholder groups.[29]

[28] UN, *Road map for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation*, May 2020

[29] IGF, *About the Internet Governance Forum*, 2023

The private sector, the technical community and civil society should be involved from the beginning when making decisions on digital topics.[30] Involving stakeholders across technical and non-technical communities, promoting inclusivity, and respecting the different cultural backgrounds are key components for designing a system approach to global AI governance. Multistakeholder engagement furthermore should meaningfully address concerns of various actors and consider power asymmetries between them.[31] This report is developed by the PNAI multi-stakeholder community through a transparent process and an open consultation.[32]

---

[30] UN, *Road map for digital cooperation: implementation of the recommendations of the High-level Panel on Digital Cooperation*, May 2020

[31] Cecilia Cabanero Verzosa and Thomas R. Fiutak, *The "How" of Multistakeholder Engagement*, ADB government brief, 2019

[32] PNAI, *PNAI Work Plan*, May 2023

4. Interoperability of AI governance

There are several approaches to regulating AI globally. The European Union's (EU) AI Act[33] to regulate development and use of AI is currently under negotiation to be adopted. China has been fast to turn proposals into rules, and for example Brazil and Canada have proposed legislation to regulate AI. Countries and regions around the world are actively making plans and pursuing their strategies to govern AI.[34] Approaches to regulate AI take different forms, for example international treaties, national legislation, regulatory sandboxes, ethical guidelines, private standards, technological solutions as well as multi-stakeholder approaches, industry self-regulation, sectoral approaches, , open source collaborations, and technological neutrality.

Interoperability is often understood as the ability of different systems to communicate and work seamlessly together. In this chapter we explore interoperability of AI governance in the global level. We argue that more emphasis should be put in analysing if and how the different initiatives to regulate and govern AI across the world could work together and through that become more impactful. As noted in the introduction, AI and its societal impact transcend boundaries of countries and regions. This chapter explores interoperability of global AI governance from a multi-stakeholder view.

We start by defining interoperability in the context of AI governance.[35] This is a critical step in clarifying the focus and scope of our multi-stakeholder writing team's work. Further, a clear definition is needed to integrate both the technical and non-technical elements to the discussion. We recognise the need to support effective cooperation and communication that is needed for building trust and a shared understanding.

Our multistakeholder group's definition of interoperability in AI governance is a framework that brings together of three key aspects: (1) the substantive tools, measures and mechanisms involved in guiding and developing AI, (2) multistakeholder interactions and interconnections, and (3) agreed ways to communicate and cooperate. All three are necessary to support a common understanding, interpretation and implementation of transborder governance of AI. The definition and our interpretation of interoperability in this context is not based on a systematic survey but was developed through consensus in our group consisting of team members from civil society, technical experts, government officials and private sector representatives. [36]

---

[33] European Parliament, *EU AI Act: first regulation on artificial intelligence*, information website, June 2023 (Accessed 10.9.2023)

[34] Harvard Business Review, *Who Is Going to Regulate AI?*, 2023

[35] There are numerous earlier definitions for interoperability. See Annex 1 for examples.

[36] Promoting interoperability doesn't necessarily mean creating identical frameworks or standardized norms for different countries and contexts. Countries can pursue a range of policy needs in AI governance counting on multistakeholder participation.

Figure 3: Defining interoperability of AI governance
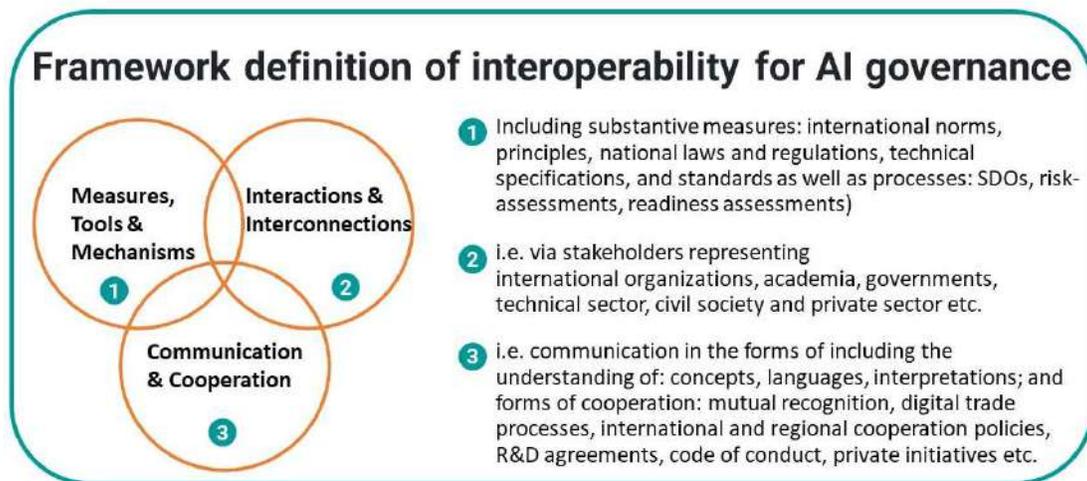
When it comes to the interoperability of AI governance, there are a number of challenges we need to face and address, including: **Ethical and regulatory Considerations** – Interoperability in AI governance raises ethical concerns related to bias, fairness, transparency, and accountability. Different jurisdictions may have varying regulations that impact data sharing, privacy, and algorithmic behavior. **Standardization** – The absence of universally accepted standards, principles, and norms for addressing interoperability in AI governance complicates efforts to create a cohesive framework.   **Semantic Interoperability** – Beyond technical compatibility, AI systems must also achieve semantic interoperability. This involves a shared understanding of the meaning, intent, nuances, and context of data and actions.

We live in a globally interconnected world and to unlock the full potential of AI we need increased interoperability in global AI governance. Ensuring harmonious coexistence among AI systems has the potential to revolutionize industries, enhance human capabilities, and drive innovation towards a more connected future.

4.1. Existing policy measures

In this section, we provide examples of international policies that have addressed interoperability of AI governance. We then move on to describe interoperability policies and initiatives in the Global South to identify and understand the commonalities and differences in technology development, as well as AI governance capacity between the South-South and North-South.

Our aim is to illustrate the most prevalent types of policies, practices, and issues. The examples we present are not exhaustive, and this report chapter is not meant to provide a full picture of the current situation globally. In this report chapter, we seek to go beyond the most cited examples of national and regional activities in governing and regulating AI, and draw from the wide expertise in our multistakeholder group to highlight regions and countries at

different stages on their AI paths.[37] We recognize there are disparities between different regions and countries in AI development and deployment, as well as different institutional and normative frameworks in place which may result in different regulatory priorities and approaches. Also the capacities of governments in regulation may diverge significantly.[38]

4.1.1. Examples of international initiatives supporting interoperability of AI governance

United Nations Secretary-General. The latest developments on the level of the United Nations were voiced in July 2023, when the United Nations Secretary-General supported proposals to establish an international agency akin to the International Atomic Energy Agency for AI in his remarks to the UN Security Council[39]. The organization would for example establish mechanisms of monitoring and governing AI. In August 2023 the first steps were taken to establish a UN high-level AI advisory body on AI. The group will undertake analysis and advance recommendations for the international governance of AI. This could include ways to ensure AI development and governance is in line with human rights, the rule of law, and the common good. The group is expected to report back on the options for global AI governance by the end of 2023.[40]

UNESCO's Recommendations on the Ethics of AI emphasize multistakeholder and adaptive governance. They call for the adoption of open standards and interoperability to facilitate collaboration and meaningful participation by marginalized groups, communities, and individuals. Moreover, the recommendations champion the rule of law as the underlying principle of AI governance[41] and promote capacity-building for judicial operators[42] and civil servants[43]. Coupled with funds and equal regional participation, this capacitation is essential to empower multiple stakeholders to coordinate among themselves and engage in meaningful debate. Ultimately, these measures can enhance cooperation, the uptake of interoperable governance tools and the legitimacy of governance fora including standard development organizations, industry fora, national legislatives and international organizations.

The **OECD's** AI Principles[44], adopted in May 2019, were designed as flexible, future-proof and values-based to facilitate interoperability on AI governance. They provide guidance for the design, deployment and use of AI systems, and for governments in shaping their AI policies. The OECD is working on promoting interoperability of AI risk management frameworks by

---

[38] Despite strong regulatory action in some regions, the institutional shape and capacities of governments diverge significantly. See: UNESCO, *Readiness assessment methodology: a tool of the Recommendation on the Ethics of Artificial Intelligence*, 2023

[39] UN, *Secretary-General's remarks to the Security Council on Artificial Intelligence*, July 2023

[40] UN, *High-Level Advisory Body on Artificial Intelligence* information webpage, 2023 (Accessed 10.9.2023)

[41] Paragraphs 47, 60 and 63: UNESCO, *Recommendation on the Ethics of Artificial Intelligence*, 2022

[42] UNESCO, *AI and the Rule of Law*, 2023

[43] UNESCO, *Digital Capacity Building for Governments*, 2023

[44] OECD, *Recommendation of the Council on Artificial Intelligence* OECD/LEGAL/0449, May 2019

identifying common guideposts to assess AI risk and impact for Trustworthy AI.[45] The goal is to help implement effective and accountable trustworthy AI systems by promoting global consistency through mapping existing and developing core standards, frameworks and guidelines for AI design to the OECD's top-level interoperability framework.[46]

G7 countries recognise the importance of interoperability in building trust in digital economies, creating open and enabling environments for responsible AI innovation. The countries support inclusive stakeholder participation in international standards. G7 countries aim to raise awareness and strengthen capacity building among stakeholders who participate in international AI technical standards development efforts and encourage adoption of international AI standards. G7 intend to collaborate with international organizations and initiatives (including the OECD, the Global Partnership on Artificial Intelligence GPAI, and UNESCO) and enhance engagement with developing and emerging economies to adopt and implement the OECD AI Principles to reinforce human rights and fundamental freedoms. They encourage collective efforts to promote interoperability between AI governance frameworks around the world for supporting AI innovation globally.[47]

Notable examples of established regional initiatives with international impact include:

At the **Council of Europe** (CoE), its Committee on Artificial Intelligence (CAI) is drafting a legally binding instrument for the development, design and application of AI systems based on the CoE's standards for human rights, democracy and the rule of law[48], and to promote innovation[49]. The latest draft version of the document was published in July 2023 to serve as the basis for further negotiations of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law.[50]

The EU is finalizing its AI Act and has also been active in international dialogue with partners outside the EU. An example of a cooperation mechanism is **the EU-US Trade and Technology Council** (TTC). It was established in 2021 as a transatlantic forum to foster cooperation in trade and technology.[51] TTC's Joint Roadmap for Trustworthy AI and Risk Management published in 2022 establishes tools, methodologies, and approaches to promote its trustworthy use of AI to support democratic values and human rights. TTC pledges to advance shared and interoperable terminologies and taxonomies (for example developing interoperable definitions of 'trustworthy' or 'bias'), cooperate in developing international technical standards and tools for trustworthy AI and risk management, and share knowledge

---

[45] OECD.AI, *Expert Group on AI Risk & Accountability* information webpage (Accessed 2.10.2023)

[46] OECD, *Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI*, OECD Digital Economy Papers, 2023

[47] G7, *G7 Digital and Tech Track Annex 5 G7 Action Plan for promoting global interoperability between tools for trustworthy AI*, 2023

[48] CoE, *CAI - Committee on Artificial Intelligence* information webpage (Accessed 19.9.2023)

[49] CoE, *The Council of Europe & Artificial Intelligence*, March 2023

[50] CoE CAI, *Consolidated working draft of the framework convention on artificial intelligence, human rights, democracy and the rule of law*, July 2023

[51] European Commission, *Digital in the EU-US Trade and Technology Council* information webpage, May 2023 (Accessed 10.9.2023)

in monitoring and measuring existing and emerging AI risks. In addition to the TTC, a successful EU interoperability framework is the European Digital Innovation Hubs (EDIH) network[52]. It is built up bottom-up driven by regional stakeholders and provides digital support and interoperability in governance (not only technologically) of services throughout the EU. Federation of services via marketplaces are becoming the *de facto* standards for AI services. This way, cooperation frameworks in the EU are shaping indeed interoperability in future AI services and applications.[53]

4.1.2. Policies and projects in the Global South on interoperability of AI governance

China. In recent years, China has conducted explorations and practices in the field of AI governance. It has established a policy framework and regulations covering R&D and ethics including the principles of "harmony and friendship, fairness and justice, inclusive sharing, respect for privacy, security and control, shared responsibility, open collaboration, and agile governance" of AI development.[54] The China academy of information and communications technology (CAICT) drafted the "Self-discipline Convention for AI" based on analysis of domestic and foreign AI ethics, laws, and strategies.[55] The Chinese government's position paper on Strengthening Ethical Governance of Artificial Intelligence[56] advocates the concepts of "people-oriented" and "intelligence for good" and ensuring all countries shared benefits of AI.[57] China encourages transnational, multi-cultural as well as multi-disciplinary exchanges and collaboration in AI, and promote participation of all countries in the major international AI ethics discussion and in international rule-making. China sees that governments should strengthen the ethical supervision of international cooperative research in AI. The position paper further calls for an international agreement on the ethics of AI to be formed on the basis of universal participation, and for the formulating of international AI governance framework,

---

[52] European Commission, *European Digital Innovation Hubs* information webpage (Accessed 12.9.2023)

[53] Regarding AI governance and interoperability, not only the AI Act, but also the Data Act and the Digital Services Act are top level regulations for the aforementioned cooperation models, both fed by public private partnerships, organizations and institutions across Europe.

[54] National Governance Specialist Committee for the New Generation Artificial Intelligence, *Governance Principles for the New Generation Artificial Intelligence--Developing Responsible Artificial Intelligence*, 2019; The Standardization Administration of China, et al, *Guidelines for the Construction of a National New Generation Artificial Intelligence Standards System*, 2020; National Governance Specialist Committee for the New Generation Artificial Intelligence, *New Generation of Artificial Intelligence Ethics Code*, 2021; National Artificial Intelligence Standardization Group, *Guidelines for the Standardization of Artificial Intelligence Ethical Governance (2023 edition)*, 2023; The Ministry of Science and Technology et al, *Measures for Ethical Review of Science and Technology (Trial)*, 2023; Cyberspace Administration of China, *Regulations on Algorithm Recommendation Management of Internet Information Services*, 2021; *Regulations on Deep Synthesis Management of Internet Information Services*, 2022; and; *Interim Measures on the Management of Generative Artificial Intelligence Services*, 2023

[55] China academy of information and communications technology (CAICT), *Self-Discipline Convention for AI, 2019; Ibid, White Paper on Trustworthy Artificial Intelligence*, 2021

[56] Ministry of Foreign Affairs of China, *China's Position Paper on Strengthening Ethical Governance of Artificial Intelligence*, 2022

standards and norms with broad consensus while fully respecting the principles and practices of AI governance in different countries.

BRICS. In August 2023, the BRICS group announced their intention to establish an AI study group to monitor AI's development and progress, expand cooperation and information exchange on AI, develop an AI governance framework to ensure the safety, reliability, controllability, equality of the AI technology.[58]

India. In recent years, various government Committees, Ministries and bodies in India have released reports and white papers to regulate and standardize AI. The 2020 Indian Artificial Intelligence Stack discussion paper identifies a need to develop uniform standards, such as various interface standards and India's AI stack. The stack will be structured across all sectors[59] and use standards developed in line with internationally agreed principles to ensure a healthier and safer environment for the evolution of AI. Two white papers by NITI Aayong (think tank of the Government of India) on responsible AI highlight that the principles for responsible AI should be grounded in the nation's values and should be compatible with international standards.[60] They should ensure a flexible approach to promote innovation and be identified based on relevant social, economic, political and cultural factors. International standards may be leveraged when there are common goals. Private sector and research institutions will create frameworks for compliance with AI standards and devise cost effective compliance with AI standards.[61]

ASEAN. The 10-member Association of Southeast Asian Nations (ASEAN) has announced the development of an ASEAN Guide on AI Governance and Ethics.[62] The guide is expected to be adopted in 2024. In its 2025 Digital Masterplan[63], ASEAN focuses on the interoperability of data sharing frameworks within and beyond the region (examples include cooperating with APEC or looking in to the European GDPR). Also improving the interoperability of e-government services, especially in digital ID and e-commerce services, is a key aim of ASEAN.

Pan-Asia Initiatives. In June 2023 Singapore launched the AI Verify Foundation to harness the collective power and contributions of the global opensource community to develop AI testing tools for the responsible use of AI. The Foundation (and its more than 60 general members from across the globe) aim to foster an open-source community that will contribute to AI testing frameworks, code base, standards and best practices and create a neutral platform for open collaboration and idea-sharing on testing and governing AI. In September 2022, the Technology for Sustainable Development Goals Alliance for Asia (Tech4SDG) was

---

[58] Gigwatch, *BRICS announces formation of AI study group*, 2023

[59] AI Standardization Committee, *Indian Artificial Intelligence Stack*, 2020

[60] India.AI, *NITI Aayog launches first of two-part approach paper on responsible AI adoption*, news article, February 2021

[61] India.AI, *Responsible AI: Part 2 - Operationalizing Principles for Responsible AI*, August 2021

[62] Fanny Potkin and Panu Wongcha-um, *Exclusive: Southeast Asia to set 'guardrails' on AI with new governance code*, news article for Reuters, June 2023

[63] ASEAN, *ASEAN Digital Masterplan 2025*, 2021

established[64] as a non-profit, non-governmental international organization. It aims to build regional consensus on areas including AI ethics and tech standards through cross-sectoral exchanges and cooperations in Asia.

Africa. Globally, Africa is still catching up to many parts of the world when it comes to designing AI strategies, governance frameworks, and AI regulation.[65] Tortoise Global AI index uses "government strategy" as one of its pillars contributing to the global ranking. In the 2023 edition of the index, several African countries appear in the top 62 for government strategy, including for example South Africa, Tunisia, Morocco, Egypt, Kenya, and Nigeria.[66] In 2021, Smart Africa, an AI initiative in collaboration with the German Agency for International Cooperation (GIZ), published a blueprint for the development of AI strategies in Africa. In the same year, the African Commission on Human and Peoples Rights (ACHPR) adopted a resolution urging governments to 'work towards a comprehensive legal and ethical governance framework for AI technologies', and 'develop a regional regulatory framework that ensured that these technologies respond to the needs of the people of the continent'.[67]

In 2023, the African Union (AU) Development Agency (AUDA-NEPAD) and the AU High-Level Panel on Emerging Technologies (APET) drafted the "African Union Artificial Intelligence (AU-AI) Continental Strategy for Africa". As countries progress in AI implementation at different levels, the role of AUDA-NEPAD Agency will be to monitor the developments and ensure that all AU member states are moving towards a common goal in the AI sector. This is important as it would enable countries to pool resources, develop common frameworks and standards and share access to data. Furthermore, global cooperation will also be needed to ensure that Africa's policies and strategies are aligned with other parts of the world.[68]

Caribbean. AI adoption is low even among the large digital leaders in the region.[69] Caribbean countries are in the embryonic stage of planning AI strategies and policies, as most are concentrating their resources on data governance, including privacy and data protection. The Caribbean realizes the deployment of AI will bring efficiencies to existing industry sectors as well as new industry opportunities. The Caribbean SIDS's (Small Island Development States) are aware of the risks and harms of AI to human-rights, culture, every-day existence and industry and see that AI governance should be centered around "do no harm" principle and enhancing safety.[70] The Caribbean AI Initiative (conducted by the UNESCO Cluster Office for the Caribbean and the Broadcasting Commission of Jamaica BCJ, with the support of UNESCO IFAP) prepared a policy roadmap in 2021.[71] The Roadmap proposed that Caribbean

---

[64] Tech4SDG, *Technology for sustainable development goals alliance for Asia*, 2022

[65] Ganiu Oloruntade and Faith Omoniyi, *Where is Africa in the global conversation on regulating AI?*, 2023; ALT Advisory, *AI Governance in Africa*, September 2022

[66] Serena Cesareo and Joseph White, *The Global AI Index*, 2023

[67] Diplo, *Artificial intelligence in Africa: Continental policies and initiatives* (Accessed 19.9.2023)

[68] AUDA-NEPAD *Artificial Intelligence for Africa: Harnessing Artificial Intelligence for Africa's Socio-economic Development*, 2021

[69] Incusservice, *Incus Services State of AI in the Caribbean Survey* (Accessed 19.9.2023)

[70] UNESCO, *Caribbean AI Roadmap*, 2021

[71] Ibid

SIDs should take a multi-stakeholder regional approach to establishing regional common values and principles on AI. It is also proposed developing and executing cross-border regulations. As to interoperability, the policy roadmap recommends the forming national and regional AI Governance bodies to manage and monitor the development of standards, code of conduct, procurement, supply guidelines, and design principles. To strengthen legislation and regulations, the roadmap also proposes establishing an AI Appeal Court and Online Dispute Resolution System.

Latin America. The Latin American context related to AI governance and regulation is still fragmented and diverse, although regional discussions are widely in line with international developments and discussions on the matter. Currently, a number of countries are in the process for adopting AI regulation, following the UNESCO Recommendation on the Ethics of AI. Cooperation mechanisms were established with a key participation from UNESCO and the CAF-Development Bank of Latin America to advance the implementation of the Recommendation. The fist Latin American and Caribbean Ministerial and High Level Summit on the Ethics of Artificial Intelligence will take place in October 2023.[72]

Reaching common standards in underlying aspects of AI regulation, such as data protection and access to information, has been a challenge with national implementation being still varied. Some advances were made within the Organization of American States (OAS), with the approval of a Regional Agenda for Digital Transformation in 2022[73]. The agenda promotes transparency and accountability in the management, publication, and use of open data and digital technologies, protecting individual privacy and personal data, as well as equity and respect for human rights and inclusive growth. It also includes several commitments to the technical interoperability of information and digital systems used for the digitalization of state services.

4.2. Cooperation policies and initiatives suggestions

The previous pages were dedicated to already existing policies and showed that laws and regulations form a heterogeneous framework and an interoperability divide. In this section, we present policies and collaboration advice that could facilitate or advance interoperability in the governance of AI for the Global South. In contrast to the previous sub-chapter which looked at each country individually, in the following pages we group together countries and regions that face similar challenges. Cooperation is a key component of interoperability. In addition to state policies and regulations, bottom-up and collaborative initiatives may become de facto standards and regulation boosters.

Internationally, we observe a wave of development in interoperability frameworks of AI governance. This is mainly driven by the OECD, UN, regional alliances as well as Standard

---

[72] UNESCO, *Chile will host the First Latin American and Caribbean Ministerial and High Level Summit on the Ethics of Artificial Intelligence*, October 2023

[73] OAS, *Regional Agenda for Digital Transformation*, June 2022

Developing Organizations. Multistakeholder cooperation and collaboration, capacity building, and adopting international standards are the key elements endorsed by them for enhancing interoperability. An emphasis is placed on convergence at the level of principles and legal instruments but on divergence in approaches of regulation in AI interoperability. Other divergences include scope of cooperation and the role given to industry or private sector. The proposals we assessed range from creating new independent international bodies specifically dedicated to AI to advocating for more regional or bilateral partnerships. Or the establishment of a global observatory[74], incorporating a blend of existing resources and frameworks with fresh initiatives. UN proposed a global AI watchdog in July 2023.[75] While private sector and industry expertise and insights are invaluable, we need to remain cautious in allowing them to dominate norm-setting or enforcement processes to prevent potential regulatory capture.

Latin America and Caribbean countries. To assess the scenario of the interoperability of AI governance in Latin America, it is important to consider that the discussion on regulating AI is in its early stages. Influence from international advances, mainly of the UNESCO's Recommendations and the EU AI and Data Acts, may impact future regulation in the region. Advances are still to be made within the OAS and the meeting of high-level authorities in Chile in October 2023, which could be the beginning of a regional initiative to determine the future of AI governance in the region and to foster further regional collaboration. High-level conversations lack wider civil society participation and multi-stakeholder perspectives, although the Latin American civil society has been active in generating evidence and recommendations to guide policy discussions regarding AI development and deployment with key concerns being the human rights impacts they have particularly in historically marginalized groups

In **Africa**, there are initial approaches to regional and global cooperation and joint strategies, but concrete results have yet to be achieved. It will be interesting to see how the AU member states will review and validate the current draft of the African Union Artificial Intelligence Continental Strategy for Africa. A continentally adopted version is expected to be launched at the AU Summit of African Heads of State and Government in January 2024.[76]

China, India and Asia. Technologically capable countries in Asia are relatively advanced in AI policymaking, initiating bottom-up cross-country and cross-sector R&D projects, as well as participating in international standard settings. Like their African counterparts, they also call for balancing the individual countries' domestic practices, values, and principles in the building of consensus on interoperability of AI governance at the global level. Compared with the hard law approach, the R&D research exchanges, or collaborations in the development of self-regulatory code (soft law) seem to be more flexible and feasible in facilitating and advancing the interoperability of global AI governance for those countries. However, the

---

[74] Carnegie Council, *The Case for a Global AI Observatory (GAIO)*, July 2023
[75] The New York Times, *U.N. Officials Urge Regulation of Artificial Intelligence*, 2023
[76] AUDA-NEPAD, *Artificial Intelligence is at the core of discussions in Rwanda as the AU High-Level Panel on Emerging Technologies convenes experts to draft the AU-AI Continental Strategy*, 2023

various rapid developments of global policies and initiatives in interoperability means that they need to strengthen their participation in regional and global regulatory AI discussions and development processes. Regulators, researchers, and enterprises should be incentivized by programmes such as funding, rewards, training etc. to participate in international regulation-making.

4.3. Recommendations on interoperability of AI governance

In the context of the continued and rapid development of generative AI, we acknowledge the strategic importance of strengthening the interoperability of AI governance and at the same time fostering a pro-innovation environment for. In addition to that, we need to avoid an "out-of-control race", in the development of AI technology itself and in the governance of AI.

Our multi-stakeholder group proposes eight steps to increase interoperability of AI governance:

- **Accurately define and agree on, what needs to be addressed on global level.** This could include already as well as emerging risks related to AI, with focus on issues that have occurred or been observed in practice. To achieve interoperability in AI governance, we propose that the development of regional and/or global regulatory policies, guidelines and principles should be agile, reflexive, and inclusive, and evolve according to the AI maturity level.

- **Encourage public and private investment in governance infrastructure.** Research institutions, NGOs and enterprises should be encouraged to conduct international research on technologies, tools that will improve security, reliability, robustness, interpretability, fairness, and accountability of AI. This can be done for example via dedicated funding programs. International funding programs with focus on interoperability of AI governance should be established.

- **Strengthen legislative cooperation.** This can be achieved using various instruments that promote international cooperation. National regulators should strengthen cross-border and pan-industry cooperation. They should ensure AI governance frameworks facilitate inclusiveness and a level playing field for all to benefit of AI. Unnecessary costs and fragmentations due to different regional requirements should be avoided as far as possible. AI legislation should always be in line with human rights principles, norms and international standards.

- **Foster regional multi-stakeholder initiatives and interlink them globally.** In this way, both regional and global cooperation will be strengthened. We need to allow different speeds of cooperation based on different levels of maturity and public policy needs. We should not lose sight of the goal of increased interoperability of AI governance.

- **Strengthen capacity building.** Providing training opportunities for stakeholders through workshops, conferences and online courses can help speed up the

knowledge-building. This is necessary to help meaningful participation in AI governance discussions.

- **Reduce regional disparities to encourage increases in maturity level.** This requires a comprehensive political, scientific and industrial exchange and cooperation. Proven best practices from regions (for example, national or regional AI strategies, research programs, industry guidelines or frameworks) should be selected with interoperability in mind for adaptability and usability for regions with lower levels of maturity.

- **Monitor and evaluate progress in reaching policy goals** set on national, regional and global levels to advance interoperability of AI governance. It is essential to continuously track progress made against goals set out earlier, identify areas requiring improvement, adjust strategies accordingly, and evaluate overall effectiveness of implemented measures.

- **Uphold and strengthen the Internet Governance Forum** (IGF) process, its regional and global multistakeholder initiatives including the Policy Network on AI. We need to foster spaces for open, transparent, inclusive and transborder consensus and capacity building of AI governance.

Figure 4: Recommendations on interoperability of AI governance

Acknowledge importance of interoperability and innovation		
Interactions and Interconnections	**Communication and Cooperation**	**Tools, Measures, and Mechanisms**
Continuous focus on definition, scope, and approach	Reduce regional disparities via comprehensive exchanges and best practices sharing	Regional and global consensus and capacity buildings
Continuous monitoring, evaluation, and action	Foster regional multistakeholder initiatives and interlink them globally	Strengthening of legislative cooperation
Encourage regional and global multistakeholder investments in AI governance infrastructure		

5. Framing AI Lifecycle for race and gender inclusion

When AI systems were first created in the 1950s, the teams behind this innovation were predominantly composed of white men. More than seventy years later, this is still often the case. In this chapter, we address this and other gender-related issues relevant to AI as well as issues of race in the context of AI. When developed and deployed responsibly, AI systems have the potential of helping to improve gender and racial equality in our societies. AI systems biases can also reinforce or generate new ways to operationalize racism, sexism, homophobia, and transphobia in society and harm marginalized groups. Race and gender are interconnected and intersect in multiple ways. We address this intersectionality and its relevance in AI context in the third part of this chapter. Gender and race are complex multifaceted concepts that encompass a wide range of identities and experiences. Gender, race, ethnical biases are often embedded in AI and data governance systems, which can lead to significant challenges for example for individuals who do not conform to traditional gender norms. The relationship between gender bias and AI should be better understood since AI technologies are too often seen as neutral.[77]

Understanding and identifying gender and race biases in AI and data governance is essential to mitigating their impact on individuals and society. Racial or gender biases in AI applications have caused harm across sectors, for example in hiring, policing, judicial sentencing, and financial decision-making.[78] We need to acknowledge biases and vulnerabilities that lead to gender biases and racial disparities experienced by people across industries and around the world.[79] To address these biases, it is necessary to take an intersectional, transdisciplinary and multistakeholder approach to ethical AI and carefully consider questions of gender, race, ethnicity, sexual orientation, regionality and socioeconomic status.[80]

Gender, race biases in AI and data governance can cause harm, but it is possible to address these issues and use AI to promote equity and inclusion. Conscious, continued, and ambitious action to debias AI systems is needed to guarantee basic human rights.[81] Data governance can play a critical role in overcoming these challenges by promoting fairness, transparency, accountability, and ethical decision-making. By increasing our efforts to address gender and race biases in AI and data governance, we can create a more equitable and just society for all.

---

[77] Sinead O'Connor and Helen Liu, *Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities*, AI & Society, May 2023

[78] John Villasenor, *Artificial intelligence and bias: Four key challenges*, commentary for Brookings, January 2019

[79] OECD.AI, *Webinar: Addressing the gender bias in artificial intelligence data*, webinar recording, March 2021 (Accessed 11.8.2023)

[80] Ibid

[81] P.S: Varsha, *How can we manage biases in artificial intelligence systems*, April 2023

5.1. Race

The UN Human Rights Council has stated that "Technology is a product of society, its values, its priorities and even its inequities, including those related to racism and intolerance".[82] The next pages of this report focus on race issues in the context of AI. The general society has low understanding of racial discrimination. There is social pressure for historically marginalized people to be responsible alone for solving the problems that are consequence of discrimination. There is also a tendency to understand technologies and technologic fields developing them as neutral and objective. These three phenomena add up to the problems we face when we want to discuss AI and ethnic-racial discriminations.

AI systems learn and reproduce what humans have taught them. Therefore, if the person or data responsible for programming/training the system is intentionally or unintentionally racist, the system will have similar tendencies. Biases, defined as "outcomes which are systematically less favorable to individuals within a particular group and where there is no relevant difference between groups that justifies such harms"[83], are inherent to AI systems. The definition of race is based on an ideology and a historical social construct used to group people. This notion was created with the aim of hierarchizing different individuals, thus creating the relational idea that if some people are superior, others are consequently inferior. Oftentimes, race divides human beings into groups based on their physical appearances, social factors, cultural backgrounds, and descent. As a social construct, race is often used by dominant groups in society to continue establishing a system of power over other categories, which is a factor that leads to racial inequalities. UN's International Convention on the Elimination of All Forms of Racial Discrimination defines racial discrimination as "any distinction, exclusion, restriction or preference based on race, color, descent, or national or ethnic origin which has the purpose or effect of nullifying or impairing the recognition, enjoyment or exercise, on an equal footing, of human rights and fundamental freedoms in the political, economic, social, cultural or any other field of public life."[84] There remains much to be done in pursuing equity and respect for people regardless of how their bodies are read and racially classified.

Humans do not receive enough quality and full-spectrum training, this means there is not enough attention paid to gender, racial and ethnic biases AI systems can develop. Most of the AI developers belong to historically privileged groups. Therefore, the persons targeted by the racial biases rarely present when AI systems are developed. This can eventually be used to their disadvantage. Including marginalized people in the development and creation processes can help underline problematic aspects that could lead to a wider full spectrum inclusion of AI systems. It is clear that we need for example more women, gender-diverse individuals, and

---

[82] Paragraph 13, United Nations Human Rights Council, *Racial discrimination and emerging digital technologies: a human rights analysis*, A/HRC/44/57, June 2020

[83] Nicol Turner Lee, Paul Resnick, and Genie Barton, *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms*, 2019

[84] Article 1(1) UN, *International Convention on the Elimination of All Forms of Racial Discrimination*, UN General Assembly resolution 2106, December 1965

Black people working in the development of AI. A diverse team can bring different perspectives and experiences to the table, which can help identify biases and create more comprehensive solutions.

Moreover, another factor leading to racial discrimination is that the AI systems we have nowadays are mostly based on how they were created years ago. Therefore, if a system created in the 1970s is based on stereotypes and certain power balances and racial inequities of that time, this is still the foundation of a system in use today. Indeed, "the problem is not surveillance technology itself, but the ways technology is deployed to reinforce pre-existing power disparities"[85]. Ideally, building blocks of such AI systems should be broken to down completely and created anew. However, this is not a feasible solution, since starting from scratch with AI systems would delay technical development. Instead, we need to focus on changing the biases that the AI systems have developed and learned from humans over the years. This would mean teaching the AI systems differently. This needs to be done rapidly, before the system understands certain biases as permanent, which would be detrimental to many persons belonging to racial and ethnic minorities.

Milner and Traub[86] define algorithmic racism as "the use of Big Data in ways that, intentionally or not, reproduce and spread racial disparities, shifting power and control away from Black and brown people and communities". This type of racism appears in many different areas of life, may it be facial recognition, medical examinations, or more basic tasks such as washing our hands, where for instance some automatic faucets do not recognize certain skin tones. This underlines the daily impact that AI systems have on people's lives. AI systems help spread stereotypes in society by, for instance, identifying people of color as janitors or criminals[87], this reinforces acts of discrimination and racism. AI should represent development and innovation; it should not bring us back to a society where minorities are excluded and repressed.

CASE: Facial recognition

Facial recognition is an instrument that has become more ubiquitous, may it be on our personal phones, but also in security mechanisms. A facial recognition software at the MIT in the USA was incapable of recognizing the faces of Black students, and they had to wear white masks on their faces to gain access to different areas of the university buildings. The team creating the software was composed of white men. They were no tests run on different skin colors to ensure that the system worked on all persons.

---

[85] Yeshimabeit Milner and Amy Traub *Data Capitalism and Algorithmic Racism*, May 2021
[86] Ibid
[87] Jeff Raikes, *AI Can Be Racist: Let's Make Sure It Works For Everyone*, article in Forbes, April 2023

A number of studies have underlined that AI systems are less capable of generating and recognizing faces of Black people.[88] For example, when AI systems are asked to recognize the gender of a face, they have an error rate of 35% for Black women, compared to an error rate of only 1% when it comes to white men.[89] This is an enormous gap. It can be explained by the fact that AI systems have not been created with the understanding of nuances of Black people's features. Many AI systems are based on stereotypical views of minorities, since the systems have not been trained well enough.[90] As a concrete example, AI systems did not recognize iconic people, such as Michelle Obama and Serena Williams, correctly.[91]

Facial and biometric recognition play an important role in police forces. Using biometric recognition in order to track and identify ethnic minorities against their will to restrict their movements and activities violates a number of human rights. Facial recognition softwares in the body cameras that police agents carry have been trained with databases of pictures including mugshot photos, in which Black people are overrepresented.[92] This leads to a disproportionality in their representation, which could translate to more arrests of Black people.[93] Big companies involved in creating facial recognition software, have shifted the blame to the people who created these databases instead of addressing the racial bias in their products.[94] AI systems have proven to be less effective in precisely recognizing the face of Black people many times. Such systems can be highly dangerous and discriminate against certain groups. Robert Williams was wrongfully arrested after a facial recognition software identified him as the thief in a burglary.[95] It was then proved that it was not him and the software had been mistaken.

AI is used to spread racist speech and incitement to discrimination, as well as violence targeted at certain groups.[96] Social media platforms have been criticized for not being able to recognize instances of racism and letting them proliferate. But is the issue really companies and developers not being able to identify such instances, or simply that they are not motivated to recognize racism in their platforms? It is problematic that the creators have not noted the problems or taken them into account when creating the platforms. Identifying racist hate

---

[88] Joy Buolamwini, *Artificial Intelligence Has a Racial and Gender Bias Problem*, article in Time magazine, February 2019

[89] Ibid

[90] Zachary Small, *Black Artists Say A.I. Shows Bias, With Algorithms Erasing Their History*, article in the New York Times, July 2023

[91] Joy Buolamwini, *Artificial Intelligence Has a Racial and Gender Bias Problem*, article in Time magazine, February 2019

[92] Sidney Perkowitz, *The Bias in the Machine: Facial Recognition Technology and Racial Disparities*, February 2021

[93] Alex Najibi, *Racial Discrimination in Face Recognition Technology*, October 2020

[94] Ibid

[95] Thaddeus L. Johnson and Natasha N. Johnson, *Police Facial Recognition Technology Can't Tell Black People Apart*, May 2023

[96] Paragraph 24, *UN Human Rights Council, Racial discrimination and emerging digital technologies: a human rights analysis, Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance*, June 2020

online requires resources that are most likely put in tasks that are considered more important in the companies' view.[97] Perhaps if companies lose profits, they might take action to tackle racist speech and incitement to discrimination in their platforms. Examples of racist speech online include Facebook's problems with white supremacist groups rallying and coordinating their actions on the platform. In 2018, the company admitted that the platform's AI systems were unable to detect problematic hate speech in certain contexts, such as this one.[98] A 2021 update stated that Meta's AI systems are responsible for detecting 97% of hate speech on Facebook, which is an improvement compared to previous years.[99] This lack of changes were not enough and a number of Black, Indigenous and people of color (BIPOC) users have reportedly quit the platform.[100]

There is an array of existing policy measures to tackle racism in AI. The United Nations' International Convention on the elimination of All Forms of Racial Discrimination states that States shall prevent racial discrimination, take effective measures where laws and policies have a discriminatory effect and conduct analyses and research to understand the causes and potential solutions to the issue of racism.[101] However, many countries fail to collect data that could help reveal the disparate impacts of emerging technologies.[102] Implementing more research projects and increasing funding could be a step closer to objective and complete analyses on the racial problem with regard to AI.

A number of policy groups and associations are combatting algorithmic racism and hate speech. One example is The Algorithmic Justice League[103] created by Dr Joy Buolamwini (the same woman who faced difficulties with MIT's facial recognition system) to raise awareness about impacts of AI on minorities, to open dialogues with researchers and policymakers, and to give a voice to the victims of AI. As stated by the Algorithmic Justice League, "we want the world to remember that who codes matters, how we code matters, and that we can code a better future".[104]

Although there are many negative examples of how companies are not addressing racial discrimination and biases in their AI systems, some have owned up to their mistakes and changed their way of doing. This is the case of X, (the company formerly known as Twitter) that removed a photo-cropping feature they had introduced on their platform, which did not

[97] Shirin Ghaffary, *The algorithms that detect hate speech online are biased against black people*, August 2019

[98] US Senate Committee on Commerce, Science and Transportation, *Facebook, Social Media Privacy, and the Use and Abuse of Data*, hearing, April 2018

[99] Mike Schroepfer, Meta, *Update on Our Progress on AI and Hate Speech Detection*, update by Meta Chief Technology Officer, February 2021

[100] Elizabeth Dwoskin, Nitasha Tiku and Craig Timberg, *Facebook's race-blind practices around hate speech came at the expense of Black users, new documents show*, article in The Washington Post, November 2021

[101] UN, *International Convention on the Elimination of All Forms of Racial Discrimination*, UN General Assembly resolution 2106, December 1965

[102] UN, *Racial discrimination and emerging digital technologies: a human rights analysis*, Report of the Special Rapporteur on contemporary forms of racism, racial discrimination, xenophobia and related intolerance A/HRC/44/57, June 2020

[103] See Algorithmic Justice League website (Accessed 1.10.2023)

[104] Ibid

recognize Black faces correctly. Twitter admitted that there was a racial bias, and therefore discrimination, which led to eliminating the feature.[105]

In terms of race, highlighting the lack of F.A.T.E (fairness, accountability, transparency, and ethics) is another approach that addresses the detection of biases more obliquely, with accountability measures designed to identify discrimination in the processing of personal data. Numerous organizations and companies as well as several researchers propose such accountability. Therefore, having the difficulties of foreseeing AI technologies outcomes as well as reverse-engineering algorithmic decisions, no single measure can be completely effective in avoiding perverse effects. Thus, where algorithmic decisions are consequential, it makes sense to combine measures that should be taken to work together. Advance measures such as F.A.T.E., combined with the retrospective checks of audits and human review of decisions, could help identify and address unfair results. A combination of these measures can complement each other and add up to more than the sum of the parts. This also would strengthen existing remedies for actionable discrimination by providing documentary evidence that could be used in litigation, creating new laws and policies, and frameworks, and developing a deeper understanding of the social implication of the different AI technologies and how we could use those results to improve them or not longer use them. Nevertheless, we need to hold companies that develop AI systems accountable for them to take appropriate measures. Without some kind of economic constraint, companies will not allocate resources and time to make the needed changes in their AI systems.

5.2. Gender

AI can be a powerful tool for women, girls and gender diverse people's empowerment. At the same time, AI can also hinder the progress towards equality if issues including representation, bias, and discrimination issues are not adequately addressed. We find that the increased development and use of AI systems, including generative AI, has magnified already existing obstacles for reaching global gender equality goals.

According to the World Health Organization (WHO), gender refers to the characteristics of women, men, girls, boys, and others that are socially constructed. This includes norms, behaviors, and roles associated with being a woman, man, girl, boy, or diverse expressions and identities, as well as relationships with each other. As a social construct, gender varies from society to society and can change over time.[106] Gender is different from sex, which refers to the different biological and physiological characteristics of females, males, and intersex persons.[107] Gender identity refers to a person's deeply felt, internal, and individual experience

---

[105] Alex Hern, *Twitter apologises for 'racist' image-cropping algorithm*, news article in The Guardian, September 2020
[106] WHO, *Gender and health*, information website (Accessed 1.10.2023)
[107] See meaning of gender in Cambridge Dictionary (Accessed 1.10.2023)

of gender, which may or may not correspond to the person's physiology or designated sex at birth.[108]

Gender biases in AI and data governance have become a concerning issue in recent years. These biases can arise in various stages, from data collection to algorithm design and decision-making.[109] The biases can be unintentional and reflect existing societal norms and stereotypes[110]. Women and gender-diverse people globally face unique challenges, which lead to underrepresentation and misrepresentation of certain groups in AI development.[111] [112] The biases can also stem from how data is collected, stored, and processed.[113] Gender biases in AI and data governance can have negative consequences, such as discrimination and unfair treatment. Data governance needs to be adequate to promote gender equality. Despite balanced datasets, gender biases still exist in AI technologies[114]. Algorithms being used need to be constantly checked for potential biases related to gender.

Studies suggest that AI can help reduce gender bias in decision-making by eliminating or minimizing the influence of biased information in the decision-making process. For instance, AI systems can be designed to exclude irrelevant information from the decision, such as a person's gender, race, or other characteristics that might introduce bias.[115] Additionally, AI can be employed to analyze large datasets and detect patterns of bias, enabling decision-makers to take measures to mitigate the impact of bias.[116] AI is already harnessed to reduce or mitigate inequalities. Examples that deserve to be celebrated include an initiative of Women's World Banking and Mujer Financiera[117] that uses machine learning to champion financial inclusion for women in Latin America and to supports women in managing their personal finances.

Absence of AI regulations in many parts of the world, the low transparency regarding AI use in different social contexts, as AI's impact on underrepresented social groups, such as women, can raise concerns. These concerns are not limited to deepening existing inequalities, but we are concerned that AI can also create new inequalities. It is clear that we need to address harmful practices in the application and development of AI. Attention needs to be paid also in situations where, AI-powered solutions help achieving notable positive results, but there still remains a need to improve the impact on historically marginalized groups, including women. A recent case and example of using automation to handle large volumes of data took

---

[108] See *meaning and definition of gender* in Merriam-Webster Dictionary (Accessed 1.10.2023)

[109] Sinead O'Connor and Helen Liu, *Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities*, AI&Society, May 2023

[110] Ibid

[111] Ibid

[112] Ibid

[113] Ardra Manasi, Subadra Panchanadeswaran, and Emily Sours, *Addressing Gender Bias to Achieve Ethical AI*, 2023

[114] Sinead O'Connor and Helen Liu, *Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities*, AI&Society, May 2023

[115] Brian Uzzi, *A Simple Tactic That Could Help Reduce Bias in AI*, article in Harvard Business Review, November 2020

[116] Emilia Chiscop-Head, *Can Artificial Intelligence reduce bias in human decision making?*, September 2019

[117] Women's World Banking, *Women's World Banking announces finalists for second annual Making Finance Work for Women Fintech Innovation Challenge*, press release, September 2020

place in Brazil during the Covid-19 pandemic. "Emergency Aid" is a case that sheds light on how automated processes can be a powerful tool for processing large volumes of information and making rapid decisions. The case also underscores the importance of continually improving these systems to ensure that they are fair, transparent, and capable.

Case: Automated Emergency Aid program in Brazil

Women were among those most negatively affected by the Covid-19 pandemic. The Brazilian Government created an Emergency Aid program with a special focus on women. The goal of the program was alleviating the economic and social effects of the pandemic and allowing the most vulnerable part of the population to maintain access to consumer goods, especially food. People applied for the Aid remotely through a mobile application developed by the government. To register, they were asked to provide personal information including full name, date of birth, CPF personal credit number, family composition, work conditions, and income. The algorithm made the decision to grant or deny of the benefit automatically without human involvement. Beneficiaries were selected through cross-referencing data from citizens registered in CadÚnico (Brazil's system for social program registration), and the public that registered through the app, with the program's eligibility criteria.

People trying to access and use Emergency Aid through the app faced difficulties. The created algorithm was not able to take into account recent changes in income and personal situations, as the databases it had access to were not consistent with the current situation of the people. This resulted in a significant portion of people who needed assistance not receiving the aid. Due to the lack of administrative ways to review the automated decision, the judicial system was the primary means of contestation and requesting human analysis for granting the benefit. This created a new problem due to the limitations of the State in offering legal assistance to the most vulnerable population. But despite all the limitations and access problems, the automated system made it possible for the benefit to reach a significant portion of the Brazilian population rapidly. Research suggests that (at least in relation to the initial rounds of Emergency Aid) the benefit had a positive impact on the income levels of Brazilians facing situations of great vulnerability.[118] This would not have been feasible for humans at that moment.

Although AI has evolved significantly in recent years, women around the world still have less access to education and training for digital technologies and AI specially. They are still underrepresented in AI research and development and in the boardrooms of the most influential AI companies. Recent studies have found that only 18% of authors at leading AI

---

[118] Maria Lucia Garcia, Aline F Pandolfi, Fabiola Xavier Leal and Aline F Stocco, *The COVID-19 pandemic, emergency aid and social work in Brazil*, 2021

conferences are women, and more than 80% of AI professors are men[119]. This disparity is extreme in the AI industry, for example at Facebook and Google women comprise only 15% and 10% of AI research staff in the companies[120]. In 2019, women represented only 18% of C-suite leaders in AI companies and top start-ups globally. Women's participation in key decision-making on AI is limited.[121]

AI can also perpetuate and amplify existing biases if not designed and used carefully[122]. For example, if AI systems are trained on historically biased data, they may make decisions that discriminate against certain groups[123]. Additionally, when the data used to train the AI system contains biases, such as historical discrimination against certain groups, the AI system may perpetuate and amplify these biases.[124] It is important for developers of AI systems to be aware of these potential biases and take steps to mitigate them. Additionally, we find it is important for companies, governments, civil society organizations, and multi-stakeholder initiatives to work together towards a more comprehensive view of AI fairness covering all its[125]. We need to increase collaboration between scholars from the worlds of technology, gender studies, and public policy to develop a shared language to assess and strengthen inclusion of women and gender diversity in all stages of AI lifecycle.

Finally, we need to acknowledge that while AI systems can potentially correct discrimination, the realization of this potential requires awareness, transparency, and oversight[126]. It is essential to acknowledge that pre-existing biases can affect the development and implementation of AI and data governance systems, as biased social norms and practices can introduce biases into machine learning systems through data[127].

Many governments are currently developing legislation to put in place mandatory AI audits, which need to integrate a gender perspective. The development of voluntary ethical frameworks is another way to guide the behaviors and actions in developing and using AI. Most frameworks have no safeguards which can undermine their application and oversight. Self-governance systems put in place by companies have been underpowered, including many internal human rights or ethical AI teams and bodies. Many companies continue to ignore harms their AI-powered products and services cause, or they are underinvesting in efforts to address them.

---

[119] UNESCO, *Artificial intelligence and gender equality: key findings of UNESCO's Global Dialogue*, 2020

[120] Ibid

[121] Celine Caira, Lucia Russo and Luis Aranda, *Artificially Inequitable? AI and closing the gender gap*, March 2023

[122] Anu Madgavkar, *A conversation on artificial intelligence and gender bias*, McKinsey podcast, April 2021

[123] Ibid

[124] Scott Gatzemeier, *AI Bias: Where Does It Come From and What Can We Do About It?*, June 2021

[125] Ghulam Mustafa Shoaib, *The Ethics of AI: How Can We Ensure its Responsible Use?*(Accessed 1.10.2023) ,

[126] Sinead O'Connor and Helen Liu, *Gender bias perpetuation and mitigation in AI technologies: challenges and opportunities*, 2023

[127] Sunny Shrestha auth and Sanchari, *Exploring gender biases in ML and AI academic research through systematic literature review*, October 2022

5.3. Exploring the Intersectionality of gender and race in AI

Intersectionality refers to the overlapping and interconnected nature of social identities, such as race and gender, and how they can lead to unique experiences of discrimination and bias. Intersectionality is a critical framework for improving fairness in AI by addressing the intersection of oppression, such as racism and sexism. The interconnected nature of social categories such as race, gender, and how they relate to systems of oppression and privilege, are at the heart of the notion of intersectionality.[128] However, the exclusive focus on identity categories in AI may divert attention from structural oppression that causes unfairness between subgroups. For instance, black women are oppressed because of the intersecting structure of racism and sexism, not just because they have intersecting identities of "black" and "women".[129]

Ethical AI requires taking an intersectional approach when addressing questions around gender, race, and ethnicity.[130] One area of concern is the intersectional discrimination faced by women, and gender-diverse people of color in the field of AI. For instance, AI algorithms exhibit bias when they perform better on recognizing men than women, and people with lighter skin tones than people with darker skin tones. There is an intersection of gender and race discrimination which results in lower accuracy recognizing women and gender diverse people with darker skin tones.[131] These discrepancies are problematic because they may lead to misidentification or bias against certain groups of people. Therefore, intersectionality is a critical framework to measure fairness, and AI algorithms are considered fair if probabilities of outcomes are the same or similar across different combinations of attributes such as gender and race.[132]

Considering intersectionality in the AI lifecycle processes helps ensure that AI systems are fair and just for all people. One strategy for improving diversity is to ensure a diverse group of individuals is involved in the entire AI lifecycle, from data collection to algorithm design to implementation. This includes people from various disciplines, cultures, genders, and backgrounds, as well as individuals with disabilities. Another strategy is to use inclusive design practices to ensure that AI systems are accessible to all individuals, regardless of their abilities or background. This approach involves end-users throughout the development process to ensure that the system is designed to be inclusive from the start.

---

[128] Youjin Kong, *Intersectional Fairness in AI? A Critical Analysis* (Accessed 1.10.2023)
[129] Ibid
[130] Ardra Manasi, Subadra Panchanadeswaran and Emily Sours, *Addressing Gender Bias to Achieve Ethical AI*, 2023
[131] Youjin Kong, *Intersectional Fairness in AI? A Critical Analysis* (Accessed 1.10.2023)
[132] Ibid

5.4. Recommendations on AI and gender/race

In conclusion, policy measures are still rather sparse, and we need to accelerate the progress in the fields of racial and gender discrimination in AI systems. Some feel that is that it is too late to solve the issues of gender and race, as they are too embedded in the AI systems. Changing the course is still possible, we just need to increase our efforts in erasing stereotypical views, diversifying teams working in AI and much more.

It will be hard to achieve an unbiased AI, but multistakeholder approach could offer a holistic way to understand, embody, and code the experiences of women, gender diverse, minorities, BIPOC into AI and other data-driven new technologies. Global cooperation and multistakeholder dialogue are vital in ensuring AI is a force for good also in the context of gender and race. We need to mitigate potential risks, and design paths that prioritize the well-being and security of historically marginalized groups and society at large in the age of AI.

Based on our work and discussions, here are our key statements and recommendations:

- We need to have clear policies and regulations in place that promote diversity and inclusion in AI. These policies should mandate that diversity and inclusion are taken into account during every stage of the process, and should encourage F.A.T.E. We need to promote obligatory human rights assessments of potentially risky AI applications.

- It is vital to include mainstream gender in national AI policies. We need to include time-bound gender-specific targets, allocation of resources, increased coherence to remove the multidimensional and discriminatory barriers faced by women, girls, and gender diverse people.

- We need to integrate a gender-responsive approach into the development, review and implementation of laws, policies and programs relevant for the digital age to combat new risks, gender stereotypes and bias in the fields of artificial intelligence, predictive algorithms and robotics.

- AI systems should be regularly audited to detect any biases that may have slipped through the cracks, and to ensure that they are functioning fairly and equitably.

- Any AI framework that aspires to be fair, accountable, transparent, and ethical must incorporate theories, perspectives from marginalized and underrepresented communities into all stages of AI lifecycle.

- We need to fund, support, and empower grassroots work and advocacy to foster inclusive dialogues on if and how gender, sexuality, race, ethnicity and other aspects of identity should be used in datasets and AI systems. Civil society and advocacy groups play an important role in uniting the voices of minorities to be heard by companies and institutions still promoting algorithmic racial discrimination.

- Governments should be obligated to conduct impact assessments and collect data on racial discrimination in AI. A voluntary approach is not sufficient.

- Meaningful inclusion and representation of impacted communities' representatives needs to be ensured in corporate committees. This also includes committees of national authorities or oversight bodies elected by a community or sector. Diversity and inclusion are cornerstones for developing safe and reliable AI.

- We need to step up our efforts in ensuring diversity, quality, and accuracy when building and curating datasets.

- It is vital to provide transparency, explainability, and accountability mechanisms for the whole AI lifecycle. This is especially important in the context of automated decision-making that could lead to discriminatory outcomes and harmful impacts on the fundamental and human rights of individuals.

By implementing these actions, we take one step closer to ensuring that AI systems are building a society that is fair, equitable, and accessible to all individuals.

6. Governing AI for a just twin transition

The convergence of AI, data, and environmental concern forms a dynamic nexus that holds immense promise for addressing pressing global challenges[133],[134]. Frontier technological capabilities can be leveraged to conserve and protect the environment — ultimately supporting a just green-digital transition, that fosters shared prosperity for people and the planet.

However, accelerated digital transformation (DX) creates both challenges and opportunities for the global green agenda. Digital solutions create new data-driven innovations for the common good. At the same time, accelerated DX can harm the environment[135]. The disproportional global effects of climate change[136] have led to increasing calls for a "just green transition". This refers to transitioning to an environmentally sustainable and climate-friendly economy that benefits all members of society. Intensifying datafication of societies, DX and the green transition are increasingly intertwined processes. We observe high levels of optimism that AI can be harnessed to accelerate progress towards a greener and more sustainable future, ultimately mitigating the global polycrises[137].

However, most of the hype on green-digital transition is mainly from countries in the Global North (GN), which (based on previous industrial revolutions) are better positioned to leverage positive impacts of technological disruptions[138]. These countries typically have relatively higher capabilities to combine the potential of accelerated DX with AI.[139] Many GN countries have the awareness and needed resources and prerequisites to optimize and implement data-innovations that in turn improve mitigation, adaptation, and monitoring of the triple planetary crises[140]. This GN techno-optimism often fails to capture the complexities of data-innovation ecosystems of low- and middle-income countries, most which are based in the Global South (GS). These countries are plagued by persistent and multidimensional structural inequities[141], including an AI divide[142] which will most likely hinder these countries' efforts towards successful twin transition[143].

Advances in AI, including the recent leaps made in generative AI, show significant potential for environmental conservation. One notable example is the use of generative models, such as Generative Adversarial Networks and Variational Autoencoders, for generating synthetic

[133] GPAI, *Climate Change and AI: Recommendations for Government Action*, 2021
[134] CODES, *Three Shifts for a Sustainable Planet in the Digital Age,* 2023
[135] Thanh Cong Truong, *The Impact of Digital Transformation on Environmental Sustainability*, 2022
[136] Kelly Dorkenoo, Murray Scown and Emily Boyd, *A critical review of disproportionality in loss and damage from climate change*, 2022
[137] Simon Torkington, *We're on the brink of a 'polycrisis' – how worried should we be?, article for the World Economic Forum,* 2023
[138] UNCTAD, *Technology and Innovation Report*, 2021
[139] Ibid
[140] European Commission, *A Green Deal Industrial Plan for the Net-Zero Age*, 2023
[141] Ibid
[142] Danni Yu, Hannah Rosenfeld and Abhishek Gupta, *The 'AI divide' between the Global North and the Global South, article for the World Economic Forum,* 2023
[143] Shamira Ahmed, *AI and the Circular Economy in Africa: Key considerations for a Just Transition,* 2022

data. Such data can aid in environmental research and conservation efforts[144]. Paradoxically, large and advanced models often demand significant computing power. Enormous amounts of energy are needed to train and support user queries, ultimately resulting in increased greenhouse gas emissions[145] and societal harms[146]. Natural resource consumption of AI compute infrastructure is a critical aspect to consider as AI technologies become more prevalent and powerful. Both the direct environmental impacts of developing, using and disposing of AI systems and related equipment, and the indirect costs and benefits of using AI applications should be taken into account when measuring and decreasing AI's environmental impact.[147] Other challenges connected to increased development and uptake of AI systems include ecological disruptions and human rights violations associated with the mineral value chains that supply the increased demand for hardware devices that bridge the gap between offline and online worlds[148].

Without robust data governance, AI can amplify or create intersectional inequities, particularly for the GS. Robust data governance (RDG) plays a pivotal role in shaping how environmental data is collected, stored, shared, and used for a wide range of applications such as advancing digital twin enabled innovation, informing climate-related policies, and scenario forecasting[149]. The responsible management of data is vital if we want to ensure that AI technologies are harnessed for the betterment of the environment. We need to ensure transparency, accountability, security, privacy, and foster data-innovations that support the social contract for data[150].

RDG can also play a significant role in mitigating the environmental impact of AI. RDG is crucial for collecting high-value data that is needed to assess the environmental impact of AI[151]. However, collecting data that aligns to the principles of data justice[152] is difficult because AI technologies and applications develop constantly, and many data ecosystems in the GS are inefficient. Collaboration among different stakeholders is essential to address these environmental concerns effectively.

GS countries are often characterized by diverse ecosystems, rich biodiversity, and unique environmental challenges. But GS is often not included as a critical player in the development of "consensus based" technical standards, norms, and regulation on the triple planetary crises, and increasingly for global AI governance[153]. For example, we know that open free,

---

[144] Ibid

[145] Kate Saenko, *Is generative AI bad for the environment? A computer scientist explains the carbon footprint of ChatGPT and its cousins*, 2023

[146] Ascelin Gordon, Afshin Jafari and Carl Higgs, The hidden cost of the AI boom: social and environmental exploitation, 2023

[147] OECD, *Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint*, 2022

[148] European Parliament, Conflict minerals: the bloody truth behind your smartphone, 2017

[149] Ibid

[150] World Bank, *World Development Report 2021: Data for Better Lives*, 2021

[151] OECD, *Measuring the environmental impacts of artificial intelligence compute and applications*, 2022

[152] GPAI, *Data Justice in Practice: A Guide for Impacted Communities*, 2022

[153] Shamira Ahmed, Dio Herdiawan Tobing and Mohammed Soliman, *Why the G20 Should Lead Multilateral Reform for Inclusive Responsible AI Governance for the Global South*, 2023

global, interoperable, reliable, and secure internet is a prerequisite for data free flow with trust, that supports innovations such as AI and strengthens respect for democratic values. Still, low- and middle-income countries are late internet adopters and not able to reap the benefits.

GS grapples with more environmental vulnerabilities and would benefit from innovative solutions facilitated by the digital revolution. We need to highlight the intersection of AI, data governance, and the environment for devising effective strategies that cater to local needs and challenges. Meaningful global cooperation for data[154] and the environment is needed to address these multidimensional and interdependent challenges. We need to ensure AI's net environmental impact is positive for the GS. Our report's focus lies on the Global South, but it is worthwhile to draw ideas from new initiatives on data governance being formulated in the EU. We need to discuss which elements could be adapted to formulate broader data governance framework(s) that can benefit the GS.

The purpose of this chapter is to dive in this interplay of AI, data governance, and the environment. We will present two case studies that demonstrate the importance of robust data governance (RDG) and responsible AI deployment in: (i) Food security and community resilience; and (ii) Climate disaster management. The following pages are a result of an iterative process that included open multi-stakeholder dialogue, collaboration and feedback from diverse experts.

6.1. Case studies

This chapter goes beyond mere analysis; it aspires to provide practical insights and recommendations on the technological environmental societal and governance challenges related to climate change and digitalisation. The selected cases illustrate the power of effective responsible AI governance.

6.1.1. Case study 1: Data governance and AI for food security and community resilience

Food security analysis and forecasting: A machine learning case study in southern Malawi.

Utilizing the MIRA (Measurement Indicators for Resilience Analysis) dataset collected through the United in Building and Advancing Life Expectations (UBALE) program[155] provides valuable insights into household characteristics, livelihood strategies, shocks, in predicting community-level vulnerability, and stressors in southern Malawi. Machine learning methods were used to analyze and forecast food security in southern Malawi to predict food insecurity levels and target assistance to vulnerable households.

---

[154] Steve MacFeely et al., *Towards an international data governance framework*, 2022
[155] USAID and CRS, *UBALE United in Building and Advancing Life Expectations*, n.d.

The findings reveal that location and self-reported welfare as the best predictors of food insecurity. The study highlights the importance of using data-driven modeling and machine learning to improve food security analysis and forecasting[156]. The case study reveals that AI holds a significant role in the transformation of food systems and in combatting food and nutrition insecurity[157]. Within the agricultural sector, AI can contribute in various ways. It can optimize or even automating certain human tasks like planting and harvesting and make utilization of natural resources more efficient.

However, when harnessing AI capabilities to increase food security, policy coherence and systems thinking should be used to reap potentials and mitigate risks. For example, small and medium-sized agricultural production units would need to make significant complementary investments, for instance in specialized infrastructure for collecting and transferring data. This means that beyond robust data governance, complimentary infrastructure policies and skills upgrading are crucial to support small holder farmers and stakeholders' digital capabilities and access to frontier technologies.

In the GS, there must be concerted efforts to ensure that formulating data governance frameworks for AI, and in particular for AI deployment aiming at increased food security in the face of climate change, important consideration is the affordability and accessibility of the data infrastructure and the information technology network, whereas this would imply ensuring availability of large data sets with high variability and high quality for GS contexts, and provided these data are findable, accessible, interoperable and reusable (FAIR).

Ideally, AI-enabled precision farming contributes to food security via improving yields while conserving valuable resources, such as water. The goal is to optimize amounts of fertilizer and herbicide based on level of soil nutrients, expected temperatures and wind speeds across farmland, establish ideal depths to plant seeds or required irrigation given a particular soil moisture, and handle different crops and adjust practices towards keeping carbon sequestered in the soil. Hence, data and AI are supposed to optimize harvests and minimize the input of resources, including water and fertilizer. This can only be achieved in the GS, if high-quality local data is available for analysis. It is fundamental that the creation of data ecosystems is done in an inclusive manner, set to solidify community resilience.

However, it is also fundamental to avoid that precision farming unevenly contributes to the expansion of monoculture farming practices leaving small-scale farms behind upon which multiple communities in the GS rely on, and to avoid generating new dependencies where farmers are locked into unfavourable commercial relationships with technology and services providing firms.

---

[156] Shahrzad Gholami, Erwin Knippenberg, James Campbell et al., *Food security analysis and forecasting: A machine learning case study in southern Malawi*, 2022

[157] FAO, *The State of Food Security and Nutrition in the World 2023*, 2023

6.1.2. Case study 2: Data governance and AI for climate disaster management

Sagar Vani is an intelligent AI app by the Indian government that has released mobile applications of and India Quake to spread information about earthquakes and other natural catastrophes. The two key mobile applications India Quake and Sagar Vani are intended to handle seismic events and marine safety, respectively. These applications are one of India's proactive approaches to utilizing intelligent technology for disaster management and maritime safety. The India Quake app is a cutting-edge effort in earthquake preparedness and response. This app provides real-time alerts regarding seismic activities in India and surrounding locations, ensuring that users have the information they need in time to make decisions during earthquake occurrences. Users of this program have access to real-time weather forecasts, predictions of the ocean's status, and notifications about large waves, which are crucial for conducting safe maritime operations.

The India Quake app was released by the National Centre for Seismology, which is under the Ministry of Earth Sciences, India. This software will provide the public with access to earthquake information in real-time. The Sagar Vani utilizes multiple communication channels through voice calls, mobile apps (User/Admin modules), multilingual SMS, audio advisories, social media such as Twitter and Facebook, GTS, email, fax, IVRS, radio and television transmission equipment, cloud channels, digital display boards, and digital display boards. It allows users to receive notifications when earthquakes occurred in various parts of the country. The app provides comprehensive information on the location, magnitude, and duration of the earthquake. This app's objective is to inform users about seismic activity and warn them to take safety measures in the case of an earthquake.

The intelligent apps have been created to offer the public essential information and services, particularly in the fields of maritime safety and earthquake preparedness. The India Quake apps, which concentrate on earthquake monitoring and early warning, aim to deliver several important outcomes so that users can receive messages and alerts for earthquakes in real-time. People can take urgent precautions to safeguard themselves during seismic events thanks to this early warning system that utilizes AI through historical earthquake data. It promotes preparedness and aids people in understanding the science behind earthquakes. Several other environmental data such as weather forecasts, statistics on the state of the ocean, tidal wave alerts, and other nautical data are all provided via the Sagar Vani apps.

The Sagar Vani and India Quake applications demonstrated India's commitment to using technology for the benefit of its people, the protection of its natural resources, and the incorporation of artificial intelligence in the application, the apps exemplify the critical role of data governance in harnessing AI and technology for disaster management and maritime safety. These applications rely on a vast array of data sources, demanding meticulous data collection, validation, and quality assurance. Data privacy and security are paramount, ensuring user information remains protected. Collaborative data sharing protocols and interoperability standards enable seamless communication between various agencies.

Ethical data usage is enforced to maintain trust, while transparency and accountability are crucial for user confidence. Inclusivity ensures information reaches a diverse audience. These apps showcase India's commitment to utilizing technology for public benefit, safeguarding natural resources, and effectively incorporating Ai into practical applications, all guided by robust data governance principles.

6.2. Key considerations for responsible AI use in the environmental sector

Multidimensional divides between the Global South and Global North. There are Multidimensional divides between the Global South and Global North on data governance, technical standards and norms, and resource allocation for environmental data and AI use. Promoting global digital public goods can play a pivotal role in addressing these divides and in advancing sustainable development, particularly in low- and middle-income countries. Its primary objective is to advocate for and facilitate the discovery, development, utilization, and investment in digital public goods. The significance of open-source software, open artificial intelligence, open data, free systems, and other forms of digital content that are freely accessible to the public is widely acknowledged on a global scale.

While the Sustainable Development Goals, the Sendai Framework for Disaster Risk Reduction 2015-2030, and the Paris Agreement, emphasize the importance of public engagement, improved access to information, and the availability of easily accessible and up-to-date data. The growing accessibility of open-access data and digital resources extends the potential for broader populations to derive benefits from their use, that supports a just transition[158], in domains such as disaster prevention, disaster management, and disaster risk reduction.

Lack of contextualization and enforcement of ethical AI standards to suit GS. Climate change cannot be addressed without addressing systematic injustices such as colonialism, racism, and uneven global power structures. The lack of contextualization and enforcement of ethical AI standards poses a significant challenge at the intersection of AI data governance and environmental sustainability, with implications for gender equality.

While AI holds immense potential to address environmental concerns, inadequate consideration of context-specific environmental challenges and the incorporation of gender perspectives can result in biased or incomplete solutions. Failure to recognize the gendered impacts of environmental issues, such as access to resources, can perpetuate inequalities. Additionally, without robust ethical standards, there's a risk of environmental data being misused or exploited, potentially harming marginalized communities and ecosystems. Robust data governance and responsible AI the environmental sector should prioritize a nuanced understanding of local contexts, gender disparities, and ethical principles to ensure that AI-driven solutions not only protect the environment but also promote gender equity and social justice.

---

[158] Ibid

Adopting a holistic approach that acknowledges and confronts systematic injustices that are deeply entrenched in our global systems is necessary when addressing climate change within the context of AI, data governance, and environmental sustainability necessitates a. Climate change is not just an isolated environmental issue; it is intrinsically connected to historical injustices, including colonialism and racism, which have led to uneven global power structures. To effectively combat climate change, we must recognize and address these interconnected challenges. Only by addressing these systematic injustices can we hope to create equitable and sustainable solutions to combat climate change and safeguard our environment for future generations.

6.3. Recommendations on responsible AI for a just twin transition

This chapter provided real-world examples of AI projects that have made a significant impact on climate-related efforts in the Global South. It also identified obstacles and difficulties in the adoption and integration of AI in sectors relevant to climate action, and that reflect the contextual realities of the Global South. Reformed multilateralism and collaborative efforts are needed to address the complex challenges at the intersection of AI, data governance, and the environment in the Global South.

The case studies highlight that while there are pockets of excellence in leveraging AI for the environment, there is a pressing need for the development of robust data governance and interdependent investments in human capital, digital infrastructure, increased research funding. We need to increase the efforts in supporting deployment and incorporation of responsible AI throughout the AI lifecycle. International collaboration, knowledge exchange, access to digital public goods, and coordinated funding for initiatives that utilize AI to address climate-related issues should be accelerated and fostered as they are of utmost importance.

Here are the recommendations and conclusions based on the discussions and work of the multi-stakeholder group:

Ensure a decolonial informed approach to data free flows with trust

- Develop policies that ensure equitable management and access to high value data sets and other digital public goods, to enable sharing of digital dividends and to promote effective just data value creation, particularly for AI use cases where there is a history of resource exploitation, that harms local communities.

- Promote transparency and accountability through each phase of the AI life cycle by developing practical and contextually relevant responsible AI and robust data governance frameworks.

- Reform multilateralism to dismantle the status quo and ensure international collaboration promotes, meaningful participation, fair consensus on technical, standards, norms, and agreements, and responsible resource management.

- Develop robust data governance frameworks to address data ownership, sharing, and protection issues concerning environmental Digital Public Goods. Promote open data standards and interoperability protocols for environmental data by prioritizing transparency, security, and accountability so that these standards facilitate data integration and harmonization.

- Data Free Flow with Trust principles should be applied to enable secure and seamless cross-border exchange of environmental data, promoting international collaboration and knowledge sharing. Equally important is advocating for equitable access to Environmental Digital Public Goods, encompassing high-value environmental datasets, AI models, and tools. These resources are vital for effective global efforts to tackle environmental issues.

Climate-Resilient Technology Adoption and Capacity Building

- Consider a decolonial informed approach for less predatory investments, loans, and financing mechanisms to improve the adoption of climate-resilient technologies and local expert lead data ecosystems in the Global South. A decolonial informed approach encourages contextually relevant incentives and capacity-building programs, including to leverage local AI-driven solutions.

- Support technology transfer, international collaboration for access to Recommendation on the Ethics of Artificial Intelligence climate technologies, and formal agreements to facilitate responsible sharing of environmental data while respecting privacy.

- Collaboration between governments, the private sector, and civil society organizations should be fostered to develop and maintain DPGs for environmental data, ensuring their sustainability, accessibility, and to provide incentives for data sharing (such as grants, awards, and recognition) can encourage organizations and individuals to contribute to Digital Public Goods, thereby enhancing global collaboration and knowledge exchange for the development of AI-driven solutions in environmental conservation and sustainability.

- Invest in educational and skill development programs to build local capacity in data science and AI.

Gender-Responsive Environmental Policies and Data Governance

- Formulate and implement gender-responsive climate and environmental policies, integrate considerations of intersectionality into leveraging data driven-AI solutions.

- Empower women and marginalized communities in decision-making processes related to environmental conservation, data governance, and AI, at all stages of the AI lifecycle.

- Ensure that AI and data governance diversity in development of solutions prioritize inclusivity in decision making (governing), and equitable access to environmental data.

Mitigate Environmental Risks and Invest in Sustainable Data Economy infrastructure

- Mitigate environmental risks of AI deployment through assessments, energy-efficient algorithms, and responsible AI practices.

- Ensure a just transition for data centres and other AI related infrastructure investments and to ensure sustainable digital development by adopting renewable energy, managing e-waste, and promoting circular economy principles, based on contextual realities while considering overlapping structural inequalities.

- Mandate monitoring and reporting of the environmental impact of AI and data operations, to create empirical evidence for informed policymaking and to encourage transparency and accountability.

These recommendations collectively offer a framework for public policymakers in the Global South to promote interoperable AI governance interventions, that harness the potential of AI and data technologies for sustainable digital development while addressing, historical injustices, promoting gender equity, and minimizing environmental harm.

ANNEX: Defining interoperability

What is interoperability? There are numerous definitions for interoperability, here are three examples:

- In the context of cross-border data transfers, **the OECD** defines interoperability as "the ability of various privacy regimes, or legal frameworks, to work together to facilitate. transborder data flows while ensuring the consistent protection of these data". The OECD Going Digital Policy Toolkit also describes interoperability as "a pragmatic arrangement to promote policy coherence in the context of a shifting regulatory environment and multiple privacy frameworks and data regulations (e.g. data localisation requirements)".

    OECD, Interoperability of privacy and data protection frameworks, OECD Going Digital Toolkit Policy Note, 2021

- A definition of interoperability found in **the EU**'s legal proposal covers both policy and technology; "Interoperability allows organizations to interact towards mutually beneficial goals. It involves the sharing of information and knowledge between organizations through the business processes they support, by means of exchanging data between their network and information systems. Interoperability ensures that data can be exchanged seamlessly. In times when processes are automated and digital technologies become part of public administrations, it is crucial that public administrations remain capable of communicating with each other. Interoperability is an important element to reach such a goal. This cannot be ensured solely by technical means. It needs agreements and established processes between different organizations, aligned data descriptions, laws that allow for those data exchanges and structured long-term cooperation. A high level of interoperability of public sector digital services is essential for the digital single market."

    European Commission, Proposal for a regulation of the European Parliament and of the Council laying down measures for a high level of public sector interoperability across the Union (Interoperable Europe Act) COM(2022) 720 final, 2022

- **China**: Facilitate the interoperability of AI systems. Artificial intelligence systems and their components have certain complexity, and different application scenarios involve different systems and components. The information interaction and sharing between systems and between components need to be guaranteed through interoperability. Artificial intelligence interoperability also involves the interoperability between different intelligent module products to achieve data interoperability, that is, different intelligent products need to have standardized interfaces. The standardization work ensures the application program interface, service, and data format of the artificial intelligence system, and defines interchangeable components, data, and transaction models through standard and compatible interfaces.

    National Artificial Intelligence Standardization General Group and the Expert Advisory Group, White Paper on Artificial Intelligence Standardization (2018 Edition), 2018

Why interoperability? Examples of benefits include:

- Technical interoperability enables cooperation between competing services and, thus, maximizes the value of communications networks. Consequently, "it facilitate[s] consumer choice, stimulate[s] innovation and tend[s] to lead to lower prices and improved quality of service for end-users."

  UNESCO, UNESCO's Internet universality indicators: a framework for assessing Internet development, 2019

- Interoperability facilitates multi-stakeholder cooperation, adaptive governance, and collaboration.

  UNESCO, Recommendation on the Ethics of Artificial Intelligence, paragraphs 47 and 77, 2022

- AI governance interoperability acts to avoid fragmentation.

IGF POLICY NETWORK ON ARTIFICIAL INTELLIGENCE | OCTOBER 2023