ETHIKAI - Paola Cantarini/ Thiago Avanci/Willis S. Guerra Filho

POLICY BRIEF

ABSTRACT

This proposal aims to address the following issues in an interdisciplinary and holistic approach, focused on the perspective of an inclusive, decolonial, sustainable, and democratic AI: how to reduce the environmental impact of AI, and what are the main challenges in Brazil as a Global South country regarding the protection of fundamental rights of vulnerable populations (indigenous and Afro-descendants); how to promote social justice and social inclusion by combining "design justice," "algorithmic justice," "epistemic justice," "data justice," and "environmental justice"; how an alternative AI governance model can contribute to achieving these objectives while aligning with innovation and technological development, integrating innovation with ethics ("metainnovation") and responsibility; and how the concept of "human-centered AI" and the presented framework for the protection of fundamental rights and environmental impact can contribute to adequate environmental protection by moving away from an anthropocentric approach toward a more holistic and sustainable understanding.

POLICY BRIEF
PART I – MOTIVATION, FOUNDATION, FRAMEWORK

Problem to be addressed and research evidence: how to reduce the environmental impact of AI, as every AI application affects the climate, and the main challenges in Brazil as a Global South country regarding the protection of fundamental rights of vulnerable populations, with a focus on indigenous and Afro-descendant communities; how to promote social justice and social inclusion through alternative AI governance models, using the theoretical framework of the Maiori data governance in New Zealand, the Toronto Declaration, which advocates for the inclusion of potentially affected groups in decision-making on design and review, and the Global Indigenous Data Alliance's "CARE Principles of Indigenous Data Governance" (https://www.gida-global.org/care). How can such a governance model and the proposal of "fundamental rights by design" contribute to aligning economic and technological development and innovation with the adequate protection of fundamental rights, integrating innovation with ethics and responsibility, including the enforcement of fundamental rights by courts in the AI era?

How will the concept of "life-centered AI," rather than just "human-centered AI," contribute to environmental protection by shifting from an anthropocentric to a holistic and sustainable perspective?

Research evidence: Algorithmic Impact Assessment (AIA) and a fundamental rights-based approach are recommended in various international documents, including the Directive from the Treasury Board of Canada (https://www.tbs-sct.gc.ca/pol/docWeng.aspx?id=32590); European Commission ("legal frameworks on fundamental rights"), Council of Europe ("Unboxing AI: 10 steps to protect Human Rights"), European Parliament ("Governing data and artificial intelligence for all - Models for sustainable and just data governance"); Federal Trade Commission (FTC), National Telecommunications and Information Administration (NTIA), Future of Privacy Forum, European Union Agency for Fundamental Rights (FRA), Dutch Data Protection Authority; Brazil's National Data Protection Authority (ANPD) (https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-publica-analise-preliminar-do-projeto-de-lei-no-2338-2023-que-dispoe-sobre-o-uso-da-inteligencia-artificial); Amnesty International and Access Now ("Toronto

Declaration"); EU High-Level Expert Group on AI (AI HLEG – "Ethical Guidelines"); Australian Human Rights Commission (2018 Project); UNESCO ("Recommendation on the Ethics of Artificial Intelligence" - https://en.unesco.org/artificial-intelligence/ethics). Additionally, studies highlight the vulnerability of the Afro-descendant population regarding facial recognition and predictive policing, and the greater potential for harm to fundamental rights in contexts with a documented historical past of discrimination and by communities systematically denied various rights throughout their history ("A Fuster and others, 'Predictably Unequal? The Effects of Machine Learning on Credit Markets,' 2021. 77 (1) J Finance 4; Safiya Umoja Noble, 'Algorithms of Oppression: How Search Engines Reinforce Racism,' NYU Press, 2018), as well as the environmental impact and decolonial perspective and vulnerability aspects of the indigenous population:
Climate Change 2021: The Physical Science Basis, Intergovernmental Panel on Climate Change, Working Group I (2021);
Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence, Mohamed et al. (2020);
Ars Technica (https://arstechnica.com/tech-policy/2020/06/police-arrested-wrong-man-based-on-facial-recognition-fail-aclu-says/);
National Institute of Standards and Technology (NIST) (https://learn.g2.com/ethics-of-facial-recognition);
Access Now, Amnesty International, and others (https://www.accessnow.org/wp-content/uploads/2021/06/BanBS-Portuguese.pdf);
Co-designing Māori data governance, https://tengira.waikato.ac.nz/__data/assets/pdf_file/0008/973763/Maori_Data_Governance_Model.pdf).

Why the relevance of the topics for the G20 agenda and priorities: the proposal is connected to sustainability, inclusion, decolonization, and social justice, aligning with the G20 agenda and priorities, especially with the values and goals of the Sherpa tracks: "digital economy" track (digital governance, information integrity, and AI), development track (poverty reduction), and "environmental and climate sustainability" track. It relates to the main objectives highlighted by the G20 Development Working Group during the Brazilian Presidency, namely, social inclusion and inequality reduction, corresponding to Sustainable Development Goal No. 10. It is linked to the central theme under the responsibilities and priorities of the Brazilian presidency, namely, "building a fairer world and a sustainable planet," and focuses on task forces 1, 2, and 5. It considers the policy briefs of the previous T20 task forces, such as the 2015 Agenda objectives, focusing on sustainability and resilience enhancement with a focus on intercultural digital ethics, and the priorities of the G20 Presidency of Brazil, emphasizing inequality reduction and digital inclusion. It brings an innovative, policy-oriented approach, proposing an alternative governance and a specific framework. The proposal has as its central axis the reduction of inequalities and inclusion in AI-related issues, linking to the UN Sustainable Development Goals: 1, 8, 9, 10, 12, 13, and 16. Regarding task forces 1, 2, and 5, the proposal aligns with the following values and objectives:
TASK FORCE 1: combating social, racial, and digital inequality, discrimination, promoting social and digital inclusion, including horizontal inequalities such as environmental racism, achieving Sustainable Development Goals (SDGs), sustainable environmental protection, and protection of vulnerable groups, sustainable and inclusive economic growth.

TASK FORCE 2: environmental impacts of AI, sustainable climate issues, inclusivity, social justice, and environmental justice, justice-oriented approach to ensure that the AI ecosystem and its benefits are distributed more equitably, emphasizing the necessary transition to a Low Carbon Economy and Sustainable Consumption/Production, promoting inclusive economic growth and neo-sustainable industrialization.
TASK FORCE 5: inclusive digital transformation, ensuring that technological advances contribute positively to achieving SDGs, prioritizing inclusion, ethics, and responsible governance in the digital era, ensuring inclusion, responsibility, and ethical considerations in AI governance, considering the perspective, values, and rights of indigenous and Afro-descendant populations, linking to SDGs 3, 4, 8, 9, 10, 11, 12, 13, 16, and 17.
History of the research: parts of this study come from two international events and discussion on two research groups. The first was held on September 25, 2021, when discussed the Brazilian bill on artificial intelligence #21/20. The “International Meeting: Brazil's AI Bill” objective was to carry out a international and multicultural analysis of the topic. The second event was “International Congress Between Brazil, Europe, And Latin America: Fundamental Rights And Updates On Procedure Law And Technologies”, and its porpoise was to discuss the judicial process in Brazil, in European countries and in Latin American countries, in the light of new technological challenges, especially those provided by artificial intelligence. It took place on May 6, 13, and 20, 2023. Those events were promoted by the CNPq Research Group Fundamental Rights, Environment, Society and Technology of the Catholic University of Santos (Brazil), and the Superior School of Advocacy, Nucleus Guarujá/SP (Brazil). Also supported the event the Society and Technology Study Center (CEST) (Brazil), the International Group of Artificial Intelligence (IGOAI) (Bahrein/UK), Global AI Ethics Institute (GAIEI) (Croatia) and EthikAI (Brazil). As mentioned, those topics were discussed also in 2023, in the internal reunions of 2023 of the Research Group Fundamental Rights, Environment, Society and Technologies and EthikAI internal circle.

PART II - EPISTEMOLOGICAL FOUNDATION

ALTERNATIVE AI GOVERNANCE, "FUNDAMENTAL RIGHTS BY DESIGN"

The main trends of our century are climate change and digital transformation, which must align with the preservation of social and democratic values, given the climate emergency (IPCC report - 1 Sixth Assessment Report of the IPCC, the United Nations body of the world’s leading climate scientists). Issues related to climate impacts involving the Global South are even more concerning (Global Partnership on AI (GPAI), "Climate change and AI: recommendations for government action," https://www.gpai.ai/projects/climate-change-and-ai.pdf), relating to power shifts in digital transformation, loss of agency, control, and sovereignty. It is interesting to note that the bills aiming to regulate AI in Brazil (PL 2120/2338/23) do not provide any recommendations regarding the environmental impact of AI, similar to GDPR and the EU AI ACT (first ediction), and these factors should be considered: CO2 emissions, electricity consumption, water consumption (used by data centers to prevent server overheating), causing water scarcity problems and worsening drought, as in Santiago, Chile, encouraging excessive and unsustainable consumption by targeted advertising algorithms, and the signing of contracts between big tech companies and fossil fuel companies undermining climate commitments.

The most recent version of the AI Act dated June 14, 2023, introduces some more robust considerations regarding the environmental impact of AI. However, it relies on non-mandatory provisions, presenting flexible rules and vague principles, voluntary codes of conduct, and the proposal for a green label, without addressing environmental sustainability comprehensively as an integral part of AI regulation. Environmental legislations, in turn, are also not equipped to deal with this issue. For instance, the EU Emissions Trading System (EU ETS) limits emissions for some specific sectors but does not address emissions related to AI. Similarly, EU water regulation, while dealing with sustainable management of water resources, does not address the use of water by AI for cooling AI data centers. Finally, the proposed EU Transparency Regulation also makes no mention of environmental sustainability, and existing EU environmental legislation does not cover such issues, especially the potential impact of generative AI. The same holds true for Brazilian legislation, creating a gap in this regard."

The present proposal for Alternative AI Governance is based on the research of Paola Cantarini, exposed in these books "AI and Law - fundamentals - vol. 5 - Epistemologies of the South" (Lumen Juris Direito, 2024), "AI Governance - a decolonial and inclusive approach," (Lumen Juris Direito, 2024), developed during a postdoctoral program at the University of São Paulo, at the Institute of Advanced Studies, Oscar Sala Chair, under the supervision of Virgilio Almeida, between the years 2022 and 2023. It brings a modular-procedural approach, as outlined in PL 2338/2023, through four interconnected layers involving legal regulation, technical design, compliance, ethical-social arrangements, and public policies, with human dignity as the top layer, the essential core of any fundamental right. Respecting these rights for all segments of the population is essential for a Democratic Rule of Law, from conception, i.e., from design and "compliance" instruments.

The proposal also involves a specific framework to be observed for the elaboration of AIA, focused on fundamental rights, involving new "fundamental rights by design" principles developed by Paola Cantarini, and for the socio-cultural characteristics of Brazil as a Global South country (Epistemologies of the South - Aníbal Quijano, Boaventura de Souza Santos), broadening the participation of vulnerable groups in the elaboration and review of AIA, and participation in the benefits of the data and AI ecosystem ("co-approach," "data sovereignty"). This aligns innovation with ethics and responsibility, betting on "metainnovation" instead of "permissionless innovation," emphasizing "responsibility for innovation" or "innovation forcing" (Hoffmann-Riem, Wolfgang. General Theory of Digital Law, pp. 13-14; p. 150 et seq.). Therefore, it considers greater vulnerability of Brazil as a Global South country (institutional and democratic fragility, recent colonial and dictatorial past (invasion of the National Congress/STF; "Freedom in the world 2022 - The Global Expansion of Authoritarian Rule," https://freedomhouse.org), such as the genocide in Myanmar (https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/), and the existence of vulnerable segments of the population (specific vulnerabilities), such as indigenous people – (informative self-determination, data sovereignty) and Afro-descendants (rights of the Quilombola population, "biometrics" and facial recognition in "policing" and predictive policing). In these cases, AI would violate the principles of Justice and Equity in cases of discrimination, racism, and compromise of linguistic diversity ("Getting the future right" from the "European Union Agency"), "bias," or "algorithmic nuisance" ("Information Fiduciaries in the Digital Age," https://balkin.blogspot.com/2014/03/information-fiduciaries-in-digital-age.html). Within the proposed framework, all potentially affected fundamental rights are taken into account ("Human

rights in national AI strategies Source": Bradley et al., 2020; "Getting the future right," "European Union Agency"), based on the principles of "fundamental rights by design," in line with "privacy by design" and "privacy by default" by Ann Cavoukian (Canada's ANPD) (Consideration 78, Article 25 of the GDPR, and LGPD, art. 46 "caput" and § 2nd). The prior preparation of the AIA would be like a kind of requirement for an additional burden of argumentation in favor of the affected fundamental rights since AI systems exacerbate structured power imbalances, affecting the most marginalized in society ("An EU Artificial Intelligence Act for fundamental rights. A Civil Society Statement").

Indigenous population: application of principles based on the Māori governance model - control - control over data and data ecosystems; jurisdiction - the physical and virtual storage location of data in Brazil; - responsibility: for all those who participate in the processing of personal data; - ethics: data processing must allow and reinforce community data governance. Decisions about data storage should prioritize sustainability for future generations.

Afro-descendant population: protection of the rights of the Quilombola population (Decree 11,447/2023) - the right to ensure access to land, infrastructure, and quality of life, productive inclusion, and local development, and participation in actions within the National Climate Change Policy and the National Policy on Payments for Environmental Services; guarantee of prior, free, and informed consultation on projects that may impact their traditional way of life. AI should also use an appropriate set of data and data training, and in content moderation, to avoid design flaws when there is no observance of the context ("emergent bias"). The Principles of the "Ubuntu" Philosophy and worldview should be observed: solidarity, reconciliation, equality, equity, and community. Solidarity: incentives for technology to create social cohesion. Reconciliation: empower unprotected communities through participation. Equity: demand for a reduction in inequality by technology companies in their products/services. Equality: adequate protection of fundamental rights, such as informative self-determination. Community: sharing the benefits of personal data used.

"LIFE CENTERED AI" AND “SUSTENTABILITY BY DESIGN”

The concept of "life centered AI" or "planet centered AI" (https://medium.com/dataseries/planet-centered-artificial-intelligence-7a208c91663a) is broader than "human centered AI," addressing the direct and indirect environmental impacts of AI, as essential to obtaining an AI "green seal," similar to the well-established ISO standard 14006 (2011) (https://www.iso.org). In the "life centered AI" approach, the multidimensionality of fundamental rights is considered, i.e., their individual, collective, and social dimensions, establishing the obligation to adopt mitigating measures to ensure that technology benefits humanity as a whole, as pointed out by UNESCO (https://www.unesco.org/en/artificial-intelligence/recommendation-ethics)" and in line with the "European Green Deal" (COM/2019/640), in favor of the "European ecological pact," focusing on the objectives of carbon neutrality by 2050, pointing to the role of design and prevention, taking into account the commitments of the UN 2030 Agenda for Sustainable Development.

FRAMEWORK AND PARAMETERS

By observing such a framework for AIA, it would be possible to obtain the "green seal" - "sustainability by design":

1. Ensure that cloud computing, if used, is included in carbon reporting and pricing policies;

2. Develop predictions about the use of energy, water, and other environmental impacts, and bring optimization strategies and possible use of renewable energy sources;

3. Prepare the AIA for environmental impact in advance - through an independent, multidisciplinary, and multiethnic team, with the participation of representatives of vulnerable groups, and publish the document on its website, containing the following steps:

a. Check the level and scope of direct and indirect environmental impacts;

b. Demonstrate why AI is used, pointing out its benefits and environmental and other fundamental rights harms;

c. Conduct a balancing procedure between conflicting fundamental rights;

d. Identify gaps and future improvements; e. Adopt the following damage mitigation or compensatory measures, proportionally, depending on the level, severity, and extent of the damage:

d.1. facilitate the creation of open data standards on AI-related environmental aspects and create an open platform to enable easy access and sharing of data;

d.2. contribute to financing the development of an international catalog of data relevant to the environment, and open-source models and software;

d.3. support accessible cloud storage systems for academic researchers, civil society, and small, medium enterprises, and startups;

d.4. support AI applications for the environment;

d.5. finance interdisciplinary research on innovation and environmental protection;

d.6. finance literacy programs and "reskilling" of AI, digital transformation, environmental impacts, and fundamental rights;

d. 7. Implement restrictions on AI training and consumption limits, depending on the social utility of the AI model on one hand, and its more utilitarian and exclusively economically oriented vision on the other.

ROADMAP – "FUNDAMENTAL RIGHTS BY DESIGN" PRINCIPLES – "EQUITY LABEL":

1. "Fundamental rights by design": be proactive and not reactive, preventive, not corrective - prior and mandatory preparation of AIA and adoption of measures in the design focused on protecting fundamental rights;

2. Be "beneficence, non-maleficence," respect human dignity and algorithmic justice, environmental justice, and epistemic justice and fundamental rights, act responsibly, carefully, transparently, and ethically;

3. "Fundamental rights by default" – protection of fundamental rights as a standard, automatically, from design and "compliance" documents, without requiring active intervention by third parties;

4. Total functionality – positive sum, not zero sum: align innovation with ethics and responsibility; promote innovation and adequate protection of rights;

5. Security and responsibility: protection of fundamental rights throughout the life cycle of AI application – prove the adoption of security measures and damage mitigation through reliable documentation;

6. Trust, transparency, and explainability: disclose the AIA on its website, in an easily accessible location; keep constant updating;

7. Respect for the principles of human-centered AI (human control of technology and respect for human values) + planet-centered AI (respect for the environment);

8. Prepare the AIA indicating the risks and benefits of AI, and its potential to violate fundamental rights, bring mitigation/compensation measures/identify gaps and future improvements. Observe the legitimacy of AIA development: independence/autonomy/interdisciplinarity/multi-ethnicity/epistemic diversity/lack of conflicts of interest, of the team, with active participation of representatives from vulnerable groups;

9. Plan and conduct human review of automated decisions affecting such groups;

10. Have an independent, autonomous, and legitimate "Oversight.

11. A constitutional amendment to ensure Fundamental Rights from the indiscriminately use of AI by the State, guaranteeing the right of one to know:

11.1 - and to decide whether to be judged by an artificial intelligence tool or equivalent automated process, or by a human judge";

11.2 - and to decide whether a administrative demand will be decided by an artificial intelligence tool or equivalent automated process, or by a civilian human public servant;

11.3- if he is interacting with an artificial intelligence tool or equivalent automated process