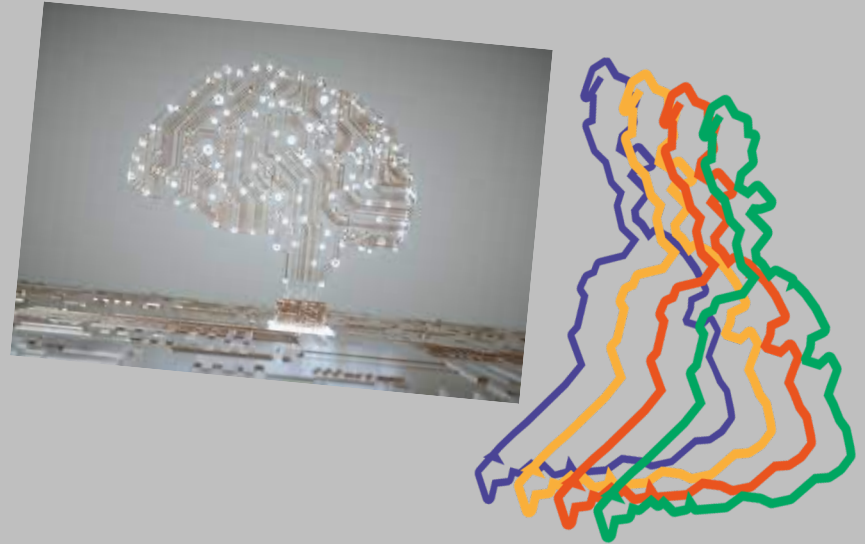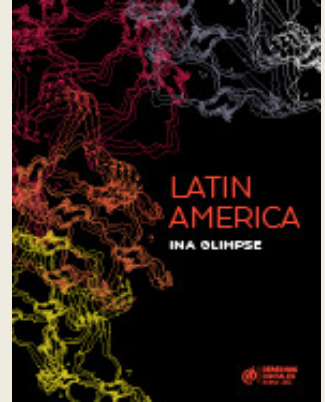# Democracy and fundamental rights: foundations of Artificial Intelligence regulation

María Paz Canales
Global Policy Advisor
Derechos Digitales

AI for what and for whom?

- Deep social and economic implications

- Impact in human development

- Therefore, touched by the State human rights obligations

Canales, María Paz. **What Do We Talk When We Talk About AI: Algorithmic Decision-Making in Latin America**. In: Latin America in Glimpse 2020", Derechos Digitales (Marzo): 19-26 https://www.derechosdigitales.org/wp-content/uploads/glimpse-2019-4-eng.pdf

So, what is expected from a Law that attempt to regulate AI?

- Principles and ethics reference are not enough

- Fundamental rights in Constitution and International HHRR obligations already provide a framework

- AI regulation should cover the concrete ways in which we ensure the promotion, respect and protection of fundamental rights

Regulation for AI as a matter of process, not technological optimism

PRINCIPLED ARTIFICIAL INTELLIGENCE

A Map of Ethical and Rights-Based Approaches to Principles for AI

Authors: Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, Madhulika Srikumar
Designers: Arushi Singh (arushisingh.net) and Melissa Axelrod (melissaaxelrod.com)

HOW TO READ:

Date, Location
Document Title
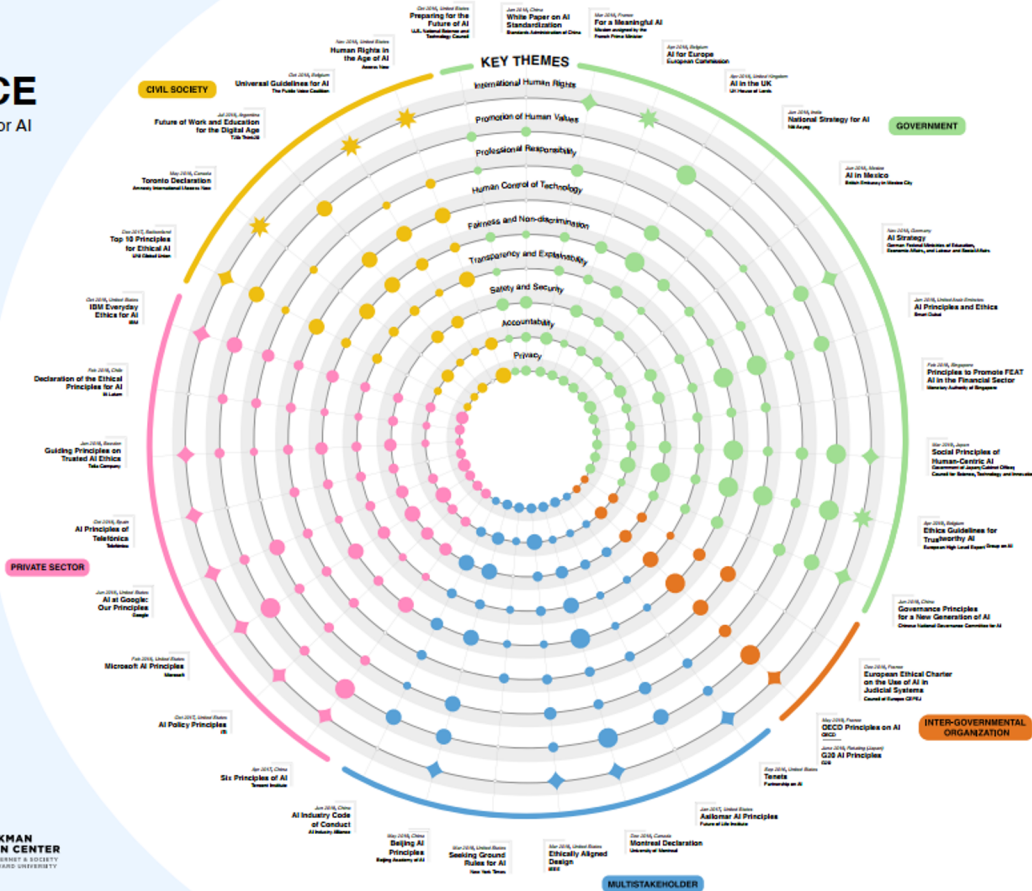Actor

COVERAGE OF THEMES:

Higher — Lower

Not referenced

◆ References International Human Rights
✦ Explicitly Adopts Human Rights Framework

The size of each dot represents the percentage of principles in that theme contained in the document. Since the number of principles per theme varies, it's informative to compare dot sizes within a theme but not between themes.

The principles within each theme are:

Privacy:
Privacy
Control over Use of Data
Consent
Privacy by Design
Recommendation for Data Protection Laws
Ability to Restrict Processing
Right to Rectification
Right to Erasure

Accountability:
Accountability
Recommendation for New Regulations
Impact Assessment
Evaluation and Auditing Requirement
Verifiability and Replicability
Liability and Legal Responsibility
Ability to Appeal
Environmental Responsibility
Creation of a Monitoring Body
Remedy for Automated Decision

Safety and Security:
Security
Safety and Reliability
Predictability
Security by Design

Transparency and Explainability:
Explainability
Transparency
Open Source Data and Algorithms
Notification when Interacting with an AI
Notification when AI Makes a Decision about an Individual
Regular Reporting Requirement
Right to Information
Open Procurement (for Government)

Fairness and Non-discrimination:
Non-discrimination and the Prevention of Bias
Fairness
Inclusiveness in Design
Inclusiveness in Impact
Representative and High Quality Data
Equality

Human Control of Technology:
Human Control of Technology
Human Review of Automated Decision
Ability to Opt out of Automated Decision

Professional Responsibility:
Multistakeholder Collaboration
Responsible Design
Consideration of Long Term Effects
Accuracy
Scientific Integrity

Promotion of Human Values:
Leveraged to Benefit Society
Human Values and Human Flourishing
Access to Technology

KEY THEMES

- International Human Rights
- Promotion of Human Values
- Professional Responsibility
- Human Control of Technology
- Fairness and Non-discrimination
- Transparency and Explainability
- Safety and Security
- Accountability
- Privacy

CIVIL SOCIETY · GOVERNMENT · PRIVATE SECTOR · INTER-GOVERNMENTAL ORGANIZATION · MULTISTAKEHOLDER

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Fjeld, Jessica, Nele Achten, Hannah Hilligoss, Adam Nagy, and Madhulika Srikumar. "Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI." Berkman Klein Center for Internet & Society, 2020. http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420
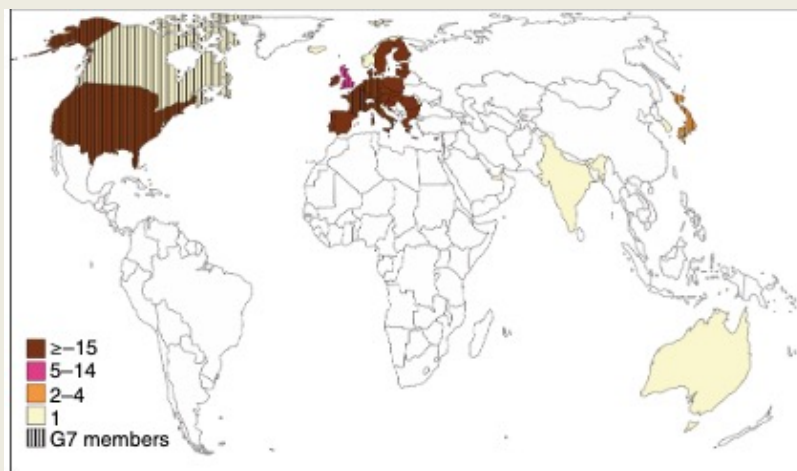
Fig. 2 | Geographic distribution of issuers of ethical AI guidelines by number of documents released. Most ethics guidelines are released in the United States (n = 21) and within the European Union (19), followed by the United Kingdom (13) and Japan (4). Canada, Iceland, Norway, the United Arab Emirates, India, Singapore, South Korea and Australia are represented

Jobin, A., Ienca, M. & Vayena, E. The global landscape of AI ethics guidelines. *Nat Mach Intell* **1,** 389–399 (2019). https://doi.org/10.1038/s42256-019-0088-2

Table 3 | Ethical principles identified in existing AI guidelines

Ethical principle	Number of documents	Included codes
Transparency	73/84	Transparency, explainability, explicability, understandability, interpretability, communication, disclosure, showing
Justice and fairness	68/84	Justice, fairness, consistency, inclusion, equality, equity, (non-)bias, (non-)discrimination, diversity, plurality, accessibility, reversibility, remedy, redress, challenge, access and distribution
Non-maleficence	60/84	Non-maleficence, security, safety, harm, protection, precaution, prevention, integrity (bodily or mental), non-subversion
Responsibility	60/84	Responsibility, accountability, liability, acting with integrity
Privacy	47/84	Privacy, personal or private information
Beneficence	41/84	Benefits, beneficence, well-being, peace, social good, common good
Freedom and autonomy	34/84	Freedom, autonomy, consent, choice, self-determination, liberty, empowerment
Trust	28/84	Trust
Sustainability	14/84	Sustainability, environment (nature), energy, resources (energy)
Dignity	13/84	Dignity
Solidarity	6/84	Solidarity, social security, cohesion

Context matter for AI a positive impact

➢ What should countries ask themself prior to implement AI solutions?

- Who is in the data?
- What are the unknown – unknown? (Data desserts)
- Internet penetration
- Availability of devices
- Digital skills in different population groups

Equity & impact in democracy of AI are more the result of the policy development for those element than chosen ethical principles or AI goals to be generally affirmed in regulation.

AI build on top of other policies

➢ AI legislation it does not work in isolation, it has to land in broader public policy directed to:

- Personal Data Protection
- Open data access
- Digital skills for the broader population
- Public Servant capacity development

Inclusive growth, sustainable development, improvement in the quality and efficiency of the services offered to the population are not the result of AI deployment in isolation.

"Acknowledges that risks to the right to privacy and other human rights can and should be minimized by adopting adequate regulation or other appropriate mechanisms, in accordance with applicable obligations under international human rights law in the conception, design, development and deployment of new and emerging digital technologies, such as artificial intelligence, by ensuring a safe, secure and high-quality data infrastructure, by exercising due diligence to assess, prevent and mitigate adverse human rights impacts, and by establishing human oversight, as well as redress mechanisms."

Human Rights Council. Resolution 48/4. Right to privacy in the digital age. Adopted October, 2021.
https://undocs.org/A/HRC/RES/48/4

What should regulation cover?

- Human Rights Impact Assessment
- Participatory design and ongoing assessment
- Equity
- Transparency
- Explainability
- Safety
- Liability
- Redress
- Independent oversight

Inteligencia artificial
e inclusión
en América Latina

https://ia.derechosdigitales.org/

"Governments should adopt a regulatory framework that sets out a procedure, particularly for public authorities, to carry out ethical impact assessments on AI systems to predict consequences, mitigate risks, avoid harmful consequences, facilitate citizen participation and address societal challenges. The assessment should also establish appropriate oversight mechanisms, including auditability, traceability and explainability, which enable the assessment of algorithms, data and design processes, as well as include external review of AI systems. Ethical impact assessments should be transparent and open to the public, where appropriate. Such assessments should also be multidisciplinary, multi-stakeholder, multicultural, pluralistic and inclusive."

Recommendation on the Ethics of Artificial Intelligence, UNESCO, adopted 23 November 2021

"Member States should ensure that harms caused through AI systems are investigated and redressed, by enacting strong enforcement mechanisms and remedial actions, to make certain that human rights and fundamental freedoms and the rule of law are respected in the digital world and in the physical world. Such mechanisms and actions should include remediation mechanisms provided by private and public sector companies. The auditability and traceability of AI systems should be promoted to this end. In addition, Member States should strengthen their institutional capacities to deliver on this commitment and should collaborate with researchers and other stakeholders to investigate, prevent and mitigate any potentially malicious uses of AI systems."

Recommendation on the Ethics of Artificial Intelligence,
UNESCO, adopted 23 November 2021

Precautionary AI deployment
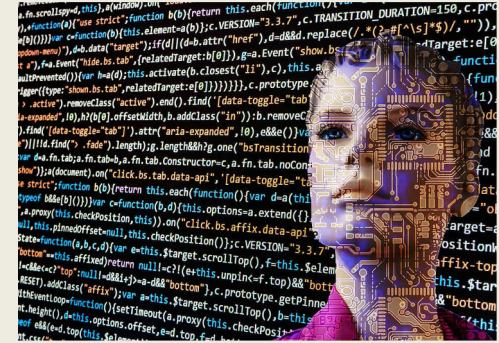
- The « guarantee of a progressive adoption of Artificial Intelligence » should not be considered as a goal by itself.

- AI adoption only should be fostered there where is evidence of its benefits for human development.

- Human oversight "when necessary" is problematic.

- Human rights impacts assessments and clear regulatory prohibitions would help to prevent harmful uses the AI.

"Expressly ban AI applications that cannot be operated in compliance with international human rights law and impose moratoriums on the sale and use of AI systems that carry a high risk for the enjoyment of human rights, unless and until adequate safeguards to protect human rights are in place.

Adopt legislative and regulatory frameworks that adequately prevent and mitigate the multifaceted adverse human rights impacts linked to the use of AI by the public and private sectors."

United Nations High Commissioner for Human Rights, Report
"The right to privacy in the digital age", September 2021

Final thoughts…

- Decision-making processes about AI adoption from governments & companies should be much more transparent for the society at large than what they currently are, often taken as an opaque administrative matter.

- The choices involved AI adoption cannot be taken in isolation of other policies & regulations that provide the contextual conditions for the AI to have an equitable impact in the population.

Final thoughts…

- There are AI implementations that can pose severe risk to fundamental rights. Human rights impacts assessments and clear regulatory prohibitions would help to prevent harmful uses the AI.

- The choices involved in AI adoption cannot be taken for rational and efficient without a transparent exercise of assessment open to the public, otherwise we will be at risk of abandoning fundamental human rights along with democratic decision-making.

Inteligencia artificial e inclusión en América Latina

https://ia.derechosdigitales.org/

mariapaz@derechosdigitales.org

@ | **DERECHOS**DIGITALES
Derechos Humanos y Tecnología en América Latina