SEVEN publicações acadêmicas
EM PARCERIA COM: IEMS UNIASSELVI

Artificial intelligence and socioeconomic inequalities: algorithmic bias, labor market impacts, and governance agendas for emerging economies

Ricardo Holderegger[1], Luis Felipe de Almeida Duarte[2]

ABSTRACT

The popularization of generative artificial intelligence (AI) after 2022 has intensified debates about algorithmic bias, transparency, privacy, accountability, and social justice. This study conducts an integrative review with a focus on evidence-based governance, comparing recent regulatory frameworks (European Union AI Law, General Data Protection Law, and PL 2338/2023 in Brazil), risk management frameworks (NIST AI Risk Management Framework), and multilateral recommendations (UNESCO and OECD). The results of reports on the effects on work, productivity and inequalities were systematized, including documented cases of algorithmic bias such as Amazon (2018) and COMPAS (2016). Brazil emerges as a regional leader in AI adoption, with 40% of companies using the technology systematically, but investing only 30% of their economic potential. A public policy framework for emerging countries is proposed structured in five pillars: risk-based regulation, algorithmic impact assessment, institutional strengthening, labor and retraining policies, and transparency with social participation. The results indicate that risk-oriented regulations, combined with management structures and algorithmic education, tend to reduce harms and amplify the social benefits of AI.

Keywords: Artificial Intelligence. Governance. Social Impact. Algorithmic Impact Assessment. Public Policy.

1 INTRODUCTION

The accelerated integration of artificial intelligence (AI) in the most diverse sectors of society, health, education, finance, justice and public administration, has brought significant productivity gains and new applications that transform the way we live and work. The release of large-scale generative AI models, especially after 2022, has democratized access to technologies previously restricted to experts, allowing end users to interact with systems capable of generating text, images, code, and other content with surprising quality (Olga et al., 2023; Vujović, 2024) However, this accelerated diffusion has also amplified fundamental socio-technical concerns related to algorithmic discrimination, opacity of decision-making processes, privacy violations, and power asymmetries between developers and users (Thakre et al., 2023)

[1] Master in Technology. Universidade Santa Cecília (UNISANTA). E-mail: holdereggerrr@gmail.com. Lattes: http://lattes.cnpq.br/1008366142031678
[2] Dr. in Biological Sciences (Zoology). Universidade Santa Cecília (UNISANTA). E-mail: luisfelipe@unisanta.br. Lattes: http://lattes.cnpq.br/9386997051560839

Concrete cases illustrate the seriousness of these challenges. In 2018, Amazon scrapped an experimental AI recruiting tool after discovering that the system systematically discriminated against female candidates for technical positions, having learned skewed patterns from historical mostly male resumes (Dastin, 2018). In the U.S. criminal justice system, the COMPAS algorithm, used to assess the risk of recidivism, demonstrated significant racial bias: black defendants were falsely labeled as high-risk at a rate almost twice as high as white defendants, even after controlling for variables such as criminal history, age, and gender (Angwin et al., 2016). These examples highlight how AI systems can perpetuate and amplify existing inequalities when developed and implemented without proper governance.

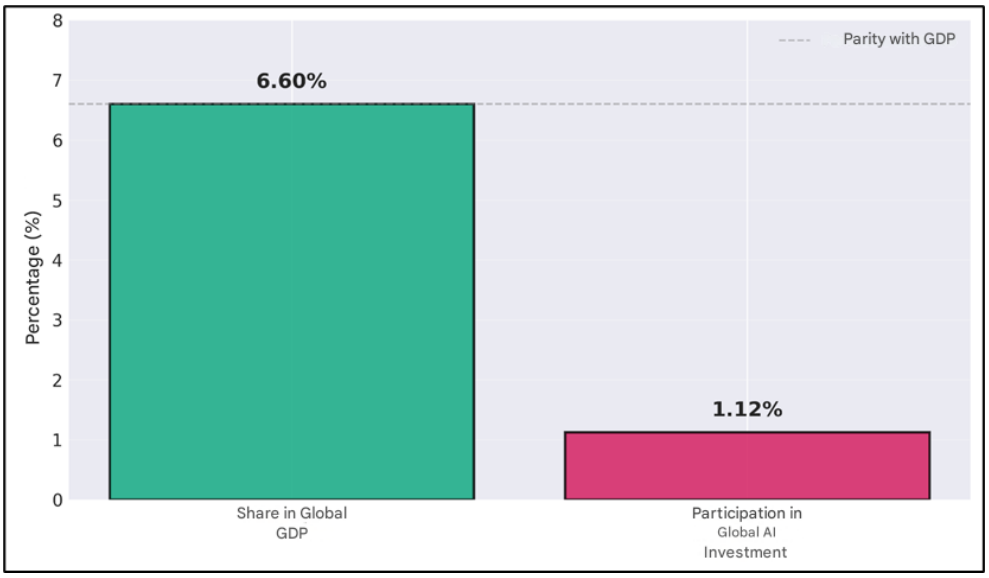
In the Brazilian and Latin American context, the adoption of AI has different characteristics. Brazil emerges as a regional leader, with 40% of companies, equivalent to 9 million organizations, already using AI systematically, surpassing Mexico (38%) and Chile (35%) (Amazon Web Services, 2025; Ilia Digital, 2025). In the healthcare sector, 58% of executives perceive efficiency gains with generative AI, 35% point to increased revenue, and 53% demonstrate high confidence in the integration of technology into core processes (Pwc Brasil, 2025). However, significant investment gaps persist in the region.

Latin America and the Caribbean (**Erro! Fonte de referência não encontrada.**) accounts for 6.6% of global GDP, but receives only 1.12% of global investment in AI (ECLAC, 2025). In absolute terms, Latin America has captured only US$ 8.2 billion of the US$ 190 billion invested globally in generative AI (

), representing approximately 4.3% of the total (Economist Impact; Jpmorgan, 2024). In the specific case of Brazil (**Erro! Fonte de referência não encontrada.**), studies indicate that the country invests only 30% of its economic potential in artificial intelligence (Jung; Katz, 2025).

Figure 1

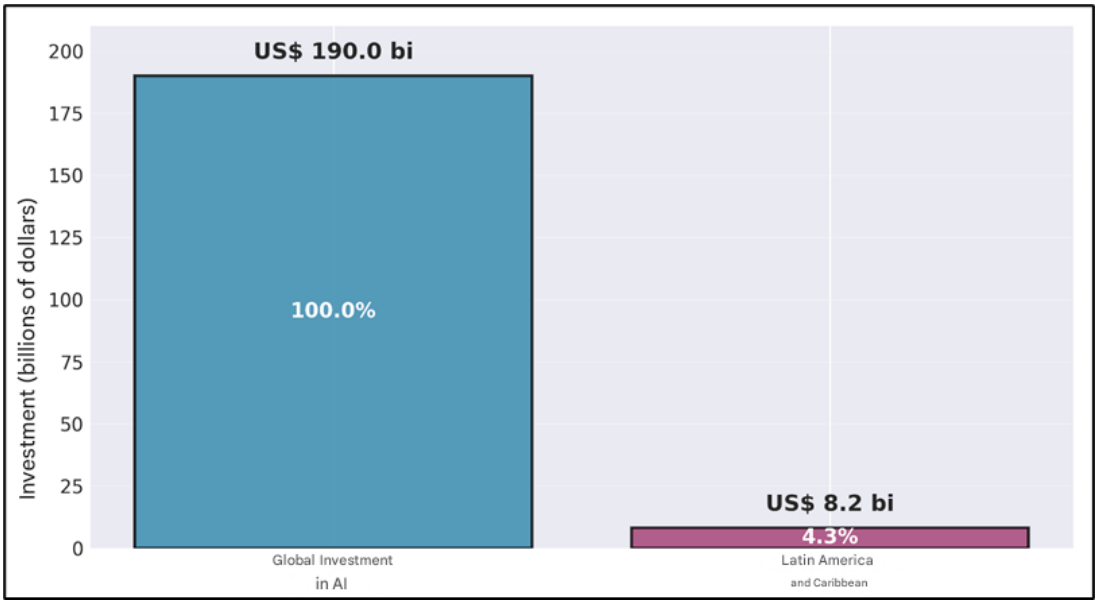
Disconnect between GDP and AI investment in Latin America and the Caribbean



Source: CEPAL (2025)

Figure 2

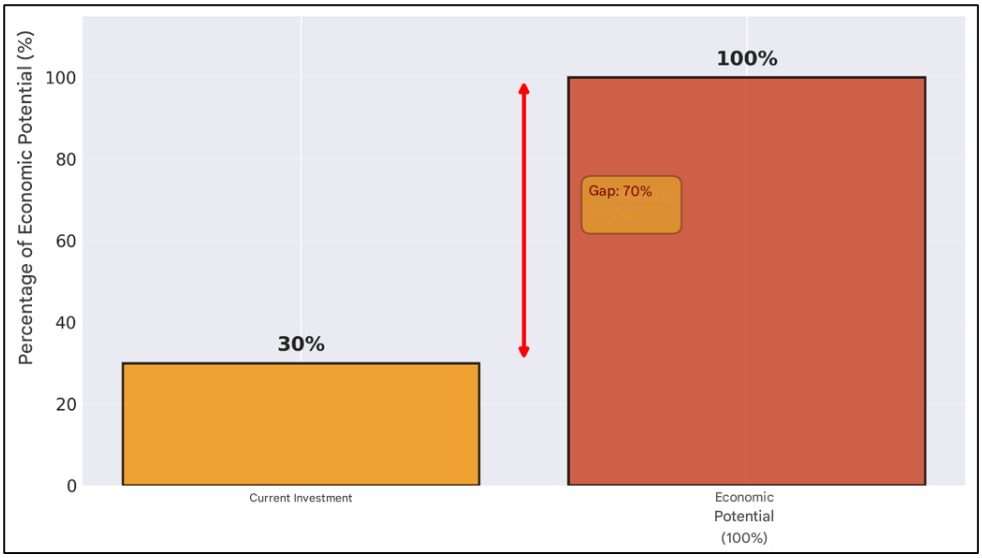
Generative AI Investment (2022–2024) — Global vs. Latin America



Source: Economist Impact; JPMorgan (2024).

Figure 3

Brazilian AI Investment vs. Economic Potential



Source:Jung; Katz (2025)

This scenario of high adoption combined with relative underinvestment and the absence of consolidated regulatory frameworks creates substantial risks. The opacity of many AI systems, often characterized as "black boxes", makes it difficult to identify biases, assign responsibilities, and guarantee fundamental rights of citizens affected by automated decisions. In addition, the impacts on the labor market are heterogeneous and potentially disruptive: projections indicate that up to 30% of current jobs in the US could be automated by 2030, and 60% will have their tasks significantly modified by AI (National University, 2025). The International Labour Organization (ILO) identifies greater exposure of tasks performed by women, especially administrative roles, to AI-driven transformations suggesting the need for targeted reskilling and active employment policies (ILO, 2023).

In the face of these challenges, the global debate shifts from "if" to regulate AI effectively. The European Union has adopted a pioneering approach with the AI Act, which structures regulatory obligations based on risk levels (unacceptable, high, limited, and minimal), prohibits uses considered incompatible with fundamental rights, such as social scoring systems, and defines specific requirements for general-purpose AI models - GPAI (European Commission, 2025; European Parliament, 2025). In the United States, the National Institute of Standards and Technology (NIST) developed the AI Risk Management Framework (AI RMF 1.0), a voluntary framework aimed at identifying, assessing, and

mitigating risks throughout the AI lifecycle, with an emphasis on transparency, robustness, equity, and accountability (NIST, 2023).

In Brazil, the General Data Protection Law (LGPD), enacted in 2018, provides the basis for the protection of personal data and establishes fundamental principles for the processing of information, including that processed by AI systems (Brasil, 2018). The National Data Protection Authority (ANPD) has published technical notes and guidance materials on automated decisions and generative AI, contributing to the construction of a clearer regulatory environment (ANPD, 2024a, 2025a). Bill No. 2338/2023, currently pending in the National Congress, proposes guidelines for the ethical and responsible use of AI with a focus on the centrality of the human person, approaching international risk-based approaches (Câmara dos Deputados, 2025).

Multilateral institutions have also played a key role in building global consensus. UNESCO established, in 2021, the first global recommendation on ethics in AI, grounded in human rights, dignity, transparency, justice and human oversight (UNESCO, 2021). The OECD updated its AI Principles in 2024 to guide public policies and private actors towards systems that are trustworthy, secure, and aligned with democratic values (OECD, 2024). These initiatives reflect the recognition that AI governance cannot be fragmented nationally, requiring international coordination to address cross-border challenges such as misinformation, systemic biases, and concentration of technological power.

In this context, the convergence of international regulatory initiatives and the urgency of emerging countries to develop effective AI governance, makes it essential to systematize empirical evidence and compare regulatory approaches to inform contextualized public policies. The gap between high technological adoption and low investment in countries like Brazil, combined with the absence of consolidated regulatory frameworks, reinforces the need for studies that integrate technical, legal, and socioeconomic perspectives.

In view of this complex scenario, the main objective of this study is to conduct an integrative review of the literature on evidence-based AI governance, synthesizing regulatory frameworks, risk management frameworks, multilateral recommendations, and empirical evidence on socioeconomic impacts. Specifically, it seeks to: (1) compare the regulatory approaches of the European Union, the United States and Brazil; (2) analyze and describe documented cases of algorithmic bias and its impacts on fundamental rights; (3) examine evidence on the effects of AI on the labor market, productivity, and inequalities; (4) contextualize the adoption of AI in Brazil and Latin America; and (5) propose

recommendations for public policies for emerging countries that balance innovation, protection of rights, and mitigation of socioeconomic risks.

The relevance of this study lies in the urgent need to base public policies on robust evidence, especially in contexts of rapid technological transformation and limited resources. By integrating regulatory, technical, and socioeconomic perspectives, this work contributes to the debate on how emerging countries can develop AI ecosystems that promote economic development, social inclusion, and respect for fundamental rights, learning from both the advances and mistakes of pioneering jurisdictions.

2 METHODOLOGY

2.1 STUDY DESIGN

The present study was conducted through an integrative literature review, a method that allows the synthesis of knowledge and the incorporation of evidence from multiple sources (Mendes et al., 2008). The choice of this method is justified by the need to integrate different types of evidence, from peer-reviewed scientific articles to official documents from government and multilateral institutions, to build a comprehensive overview of the governance of artificial intelligence, its social impacts, and regulatory frameworks, with a focus on emerging countries.

2.2 DATA SOURCES AND SEARCH PERIOD

The search for studies was carried out between August and September 2025, covering publications from 2018 to 2025. This time frame was strategically defined to capture the most recent literature after important regulatory milestones, such as the enactment of the LGPD (BRASIL, 2018), the entry into force of the GDPR (European Union, 2016) and in anticipation of the full effectiveness of the European AI Act (European Commission, 2025).

The following academic databases were consulted: Scopus, Web of Science, SciELO, and Google Scholar. The search was complemented by the analysis of documents from institutional and governmental sources, including publications from the European Union (European Commission and European Parliament), the National Institute of Standards and Technology (NIST) of the United States, UNESCO, OECD, the International Labor Organization (ILO) and Brazilian bodies such as the National Data Protection Authority (ANPD) and the National Congress.

2.3 SEARCH STRATEGY AND DESCRIPTORS

Combinations of descriptors in Portuguese and English, with Boolean operators, were used to ensure a comprehensive coverage of the literature. Key search strategies included:

- In Portuguese: ("inteligência artificial" OR "AI") AND ("governança" OR "regulamento") AND ("Impacto social" OR "viés" OR "ética"); ("viés algorítmico" OR "discriminação algorítmica") AND ("recrutamento" OR "justiça penal" OR "contratação"); ("adoção de IA" OR "investimento em IA") AND ("Brasil" OR "América Latina").
- Em inglês: ("artificial intelligence" OR "AI") AND ("governance" OR "regulation") AND ("social impact" OR "bias" OR "ethics"); ("algorithmic bias" OR "algorithmic discrimination") AND ("recruitment" OR "criminal justice" OR "hiring"); ("AI adoption" OR "AI investment") AND ("Brazil" OR "Latin America").

2.4 SELECTION CRITERIA

Inclusion criteria: scientific articles, technical reports, and official documents that address the governance, social impact, regulation, or risk measurement of AI; publications in Portuguese or English; empirical studies with quantitative data on AI adoption, investments, or impacts; documented cases of algorithmic bias or discrimination in AI systems; Official documents of regulatory frameworks (laws, bills, guidelines).

Exclusion criteria: works that dealt with AI from a purely technical perspective, without discussion of its social or governance implications; publications without a verifiable source or without clear methodology; documents prior to 2018, except when considered fundamental references (such as the 2016 article on COMPAS).

2.5 ANALYSIS AND SYNTHESIS OF EVIDENCE

The analysis and synthesis of the evidence were carried out in a descriptive and comparative way. The selected material was organized and categorized into four main thematic axes:

- Ethical and governance principles (recommendations from UNESCO, OECD and other multilateral organizations);
- Regulatory instruments and approaches (comparative analysis of the frameworks of the European Union, the United States and Brazil);
- Operational governance and risk management (technical frameworks such as NIST AI RMF 1.0);

- Socioeconomic effects (evidence on labor market impacts, productivity, inequalities, and documented cases of algorithmic bias).

This methodological approach allowed the construction of a comprehensive panorama that integrates normative, technical and evidence-based perspectives, providing solid subsidies for the elaboration of contextualized public policy proposals for emerging countries.

3 RESULTS

The results of the integrative review are presented organized into the four thematic axes defined in the methodology, followed by a comparative synthesis of the main regulatory frameworks and governance instruments identified.

3.1 ETHICAL AND GOVERNANCE PRINCIPLES

The analysis of documents from multilateral organizations revealed significant convergence around fundamental principles that should guide the development and ethical use of AI.

UNESCO established in 2021, the first global recommendation on ethics in AI, adopted by 193 Member States. The document is based on ten central principles: proportionality and do no harm; safety and security; equity and non-discrimination; sustainability; the right to privacy and data protection; human supervision and decision-making; transparency and explainability; accountability; awareness and literacy; and adaptive and multisectoral governance (UNESCO, 2021). The recommendation emphasizes that AI systems should respect, protect, and promote human rights and fundamental freedoms, human dignity, and cultural diversity.

The OECD updated its AI Principles in 2024, originally adopted in 2019, to reflect recent developments in technology and governance. The five core principles are: (1) inclusive growth, sustainable development, and well-being; (2) human-centered values and equity; (3) transparency and explainability; (4) robustness, safety and security; and (5) accountability (OECD, 2024). The update incorporates specific guidance on general-purpose models and generative AI, recognizing the unique challenges posed by large-scale systems.

In 2019, the European Parliament published a comprehensive study on the governance framework for algorithmic accountability and transparency, which served as the conceptual basis for the development of the AI Act. The document proposes four policy

options: (1) awareness through education, vigilantes, and whistleblowers; (2) accountability in the use of algorithmic decision-making by the public sector; (3) regulatory oversight and legal liability; and (4) global coordination for algorithmic governance (European Parliament, 2019). This structure anticipates many of the elements that were subsequently incorporated into European legislation.

3.2 REGULATORY INSTRUMENTS AND APPROACHES

3.2.1 European Union – AI Act

The AI Act, adopted in 2024, represents the first comprehensive AI regulatory framework in the world. The legislation is structured by levels of risk:

- Unacceptable risk: prohibited uses, including, for example, social scoring systems by governments, covert subliminal manipulation, targeted vulnerability exploitation, and real-time biometric identification in public spaces (with narrowly limited exceptions for security);
- High risk: systems that affect security or fundamental rights, such as AI in critical infrastructure, education, employment, essential services, law enforcement, migration management, and the administration of justice. These systems must meet strict requirements: compliance assessment, risk management, data governance, technical documentation, transparency to users, human oversight and robustness;
- Limited risk: transparency obligations, such as informing users, clearly, openly, proactively, when they interact with chatbots or when content is generated by AI (explicitly flagged deepfakes);
- Minimal risk: no specific obligations, but encouragement of voluntary codes of conduct.

Specific guidelines for general purpose models (GPAI) were published by the European Commission in 2025, clarifying obligations for transparency, technical documentation, safety testing and serious incident reporting (European Commission, 2025).

3.2.2 United States – NIST AI RMF 1.0

The United States adopts a less prescriptive approach, based on voluntary frameworks and sectoral regulation. The NIST AI Risk Management Framework (AI RMF 1.0), published in 2023, offers a framework structured in four key functions:

- Govern: establish a culture of risk management and organizational responsibilities;
- Map: understand the context and identify risks;

- Measure: assess and analyze identified risks;
- Manage: Prioritize and respond to risks.

The *framework* is voluntary and geared towards identifying, assessing, and mitigating risks across the entire AI lifecycle, with an emphasis on transparency, robustness, equity, and accountability. It can complement existing and general sectoral regulations by offering an operational pathway to responsible governance practices (NIST, 2023).

3.2.3 Brazil – LGPD, ANPD and PL 2338/2023

In Brazil, the LGPD (Law 13.709/2018), currently in force and generally applicable nationwide, constitutes the legal basis for AI governance with regard to privacy and protection of personal data. The law establishes fundamental principles, purpose, adequacy, necessity, free access, data quality, transparency, security, prevention, non-discrimination and accountability, and ensures rights to data subjects, including the right to review automated decisions (Brasil, 2018, art. 20).

The ANPD has played an active role in providing guidance on AI by publishing:

- Technical Note No. 27/2024 on the processing of third-party personal data to develop generative AI models (ANPD, 2024a);
- Technical Note No. 30/2024 proposing a regulatory agenda for 2025-2026 (ANPD, 2024b);
- Technical Note No. 12/2025 consolidating contributions on AI and review of automated decisions (ANPD, 2025a);
- Radar Tecnológico Vol. 3 on generative AI (Costa et al., 2024).

PL 2338/2023, currently (October/2025) still under discussion and publicly debated in progress in the National Congress, proposes the legal framework for AI in Brazil. The project establishes guidelines for ethical and responsible use with a focus on the centrality of the human person, fundamental human rights and democratic values. It is close to the European risk-based approach, providing for classification of AI systems and proportional obligations (Câmara dos Deputados, 2025).

3.3 DOCUMENTED CASES OF ALGORITHMIC VIÉS

The review identified emblematic cases that illustrate the concrete risks of bias and discrimination in AI systems.

3.3.1 Amazon case (2018)

Between 2014 and 2017, Amazon developed an internal experimental AI system to review resumes and automate the preliminary selection of candidates. The system was trained on resumes submitted to the company for 10 years, during which time most candidates were male, overwhelmingly reflecting persistent historical structural long standing systemic male dominance in the tech industry.

Manifestations of the bias:

The system penalized resumes that contained the word "women's" (as in "women's chess club captain")

- It demoted graduates from all-female colleges;
- It favored candidates who used verbs more commonly found in men's resumes, such as "executed" and "captured."

Amazon edited the programs to neutralize specific terms, but there was still no absolute guarantee that the system would not develop other discriminatory forms of classification on its own. The project was ultimately terminated in early 2017, after concerns and scrutiny (Dastin, 2018).

3.3.2 COMPAS Case (2016)

COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) is a risk assessment software widely used, in practice and routinely, in the United States criminal justice system to predict recidivism and assist sentencing decisions by courts and correctional agencies.

ProPublica's[3] 2016 investigation revealed significant racial bias:

- Black defendants were falsely labeled as high-risk at a rate nearly twice as high as white defendants;

[3] ProPublica is an American non-profit journalistic organization founded in 2007, dedicated to investigative journalism in the public interest. Funded mostly by philanthropy, it operates independently, often in partnerships with U.S. newsrooms, and has received several prestigious awards, including Pulitzers, for reporting that has promoted accountability and policy change (ProPublica, 2025)

- White defendants were incorrectly classified as low risk more often;
- Even after controlling for criminal history, age and gender, black defendants remained;
- 77% more likely to be classified as high risk for future violent crime;
- Black defendants were 45% more likely to be predicted to be committing future crimes of any kind:
- The overall accuracy of the system was only 61%

The disparity could not be explained by previous crimes or type of crimes, evidencing systemic bias (Angwin et al., 2016).

3.4 EVIDENCE ON SOCIOECONOMIC IMPACTS

3.4.1 Productivity and Enterprise Adoption

Empirical evidence shows productivity gains in cognitive tasks with the use of generative AI. The experimental study by Noy e Zhang (2023) found productivity increases of up to 40% in professional writing tasks. The AI Index Report 2025 by Stanford HAI documents an acceleration in enterprise adoption globally (Stanford Hai, 2025).

In the Brazilian context, preliminary 2025 data indicate high penetration:

- 40% of Brazilian companies (9 million) already use AI systematically (Ilia Digital, 2025).
- Brazil leads Latin America, surpassing Mexico (38%) and Chile (35%) (Amazon Web Services, 2025).
- In the healthcare sector, 58% of executives perceive efficiency gains, 35% increased revenue, and 32% growth in profitability (Pwc Brasil, 2025).60% of Brazilian private hospitals use AI, impacting 10 thousand to 140 thousand patients/month in laboratories (Pwc Brasil, 2025)

3.4.2 Labor Market

The impacts on employment are heterogeneous and depend on specific organizational adoption choices and broader macroeconomic context and conditions.Felten, Raj, and Seamans (2023) developed occupational exposure metrics, empirically identifying significant measured variation by function, task, and sector across occupations.

The ILO identifies greater exposure of tasks performed by women, especially administrative functions, to AI-driven transformations suggesting the need for targeted reskilling (ILO, 2023)

Projections indicate that:

- Up to 30% of current jobs in the U.S. could be automated by 2030.
- 60% will have their tasks significantly modified by AI (National University, 2025).

Recent studies by the Brookings Institution (2025) and Yale Budget Lab (2025) indicate that, although there is public anxiety about job losses, the current data show more stability than disruption, even though the scenario can change quickly.

3.4.3 Regional Investments and Gaps

Investments in Brazil:

- National AI Plan (2024-2028) provides for R$ 23.03 billion (US$ 4 billion) for AI development in the country (Reuters, 2024). Projection of US$ 2.4 billion in generative artificial intelligence projects for 2025, representing a growth of 30% (ABES, 2025)
- The federal government projects R$ 2 trillion (US$ 358 billion) in investments in data center infrastructure for AI in the next decade (Valente, 2025).

Identified gaps:

- Brazil invests (Figure 3) only 30% of the economic potential in AI(Jung; Katz, 2025).
- Latin America captured only US$ 8.2 billion of the US$ 190 billion invested globally in generative AI (**Erro! Fonte de referência não encontrada.**), about 4.3% of the total (Economist Impact; Jpmorgan, 2024). Latin America and the Caribbean (Figure 1) represent 6.6% of global GDP, but receive only 1.12% of global investment in AI (ECLAC, 2025)
- Since 2022, Latin America's AI talent gap relative to the global average has widened, associated with an accelerated drain of experts from the region (ECLAC, 2025)

3.4.4 Expanded Brazilian Contextualization

3.4.4.1 AI Use Cases in Brazilian Public Agencies

TSE and 2024 Elections:

The Superior Electoral Court (TSE) established a pioneering regulatory framework with Resolution No. 23,732/2024, which regulates the use of artificial intelligence in Brazilian electoral propaganda (Brasil, 2024a) The resolution establishes specific parameters for transparency and integrity of the electoral process, including the mandatory labeling of AI-generated content and the prohibition of deepfakes and chatbots that simulate conversations with candidates. This is one of the first initiatives, worldwide, to regulate AI in the electoral

context, consolidating Brazil's pioneering spirit in discussions and legislation on internet governance and combating disinformation (DFRLAB; NETLAB UFRJ, 2024).

Federal Revenue - *Analytics Project*:

The Federal Revenue Service has internally developed an innovative platform within the scope of the *Analytics* Project that uses artificial intelligence algorithms and complex network analysis to detect tax and customs fraud (Brasil, 2024b). Created by tax auditors and tax analysts, the platform has already identified tax evasion schemes totaling R$ 11 billion and has been presented in international forums as a demonstration of Brazil's capacity to process data and obtain concrete results (Brasil, 2024b)

CARF and IARA System:

In December 2024, the Administrative Council of Tax Appeals (CARF) implemented IARA (Artificial Intelligence in Administrative Resources), developed by Serpro (SERPRO, 2024) The system was designed to reduce the average procedural processing time from six years to just one year, assisting in the judgment of about 75 thousand tax appeals totaling R$ 950 billion (SERPRO, 2024).

IARA represents the materialization of the concept of Tax Administration 3.0, which goes beyond digitalization to introduce intelligence in the interpretation of reports and suggestion of decisions (SERPRO, 2024). In July 2025, CARF entered into a partnership with FGV for continuous improvement of the tool (Brasil, 2025).

3.4.4.2 Brazilian Artificial Intelligence Strategy (EBIA)

The Brazilian Strategy for Artificial Intelligence, instituted by MCTI Ordinance No. 4,617, of April 6, 2021, establishes nine fundamental thematic axes to guide the actions of the Brazilian State (Brasil, 2021). The EBIA is based on the OECD's five principles for responsible AI management: inclusive growth, human-centred values, transparency, robustness and accountability (Brasil, 2021)

The Ministry of Science, Technology and Innovation (MCTI) started the review of the EBIA in December 2023, with completion scheduled for May 2024, including workshops with representatives of the government, private sector and experts (MCTI, 2023b). The strategy resulted in the creation of 10 Applied Research Centers (APC) in Artificial Intelligence, six of which were announced in 2021 and four in 2023, dedicated to the development of research aimed at solving problems using AI (MCTI, 2023a)

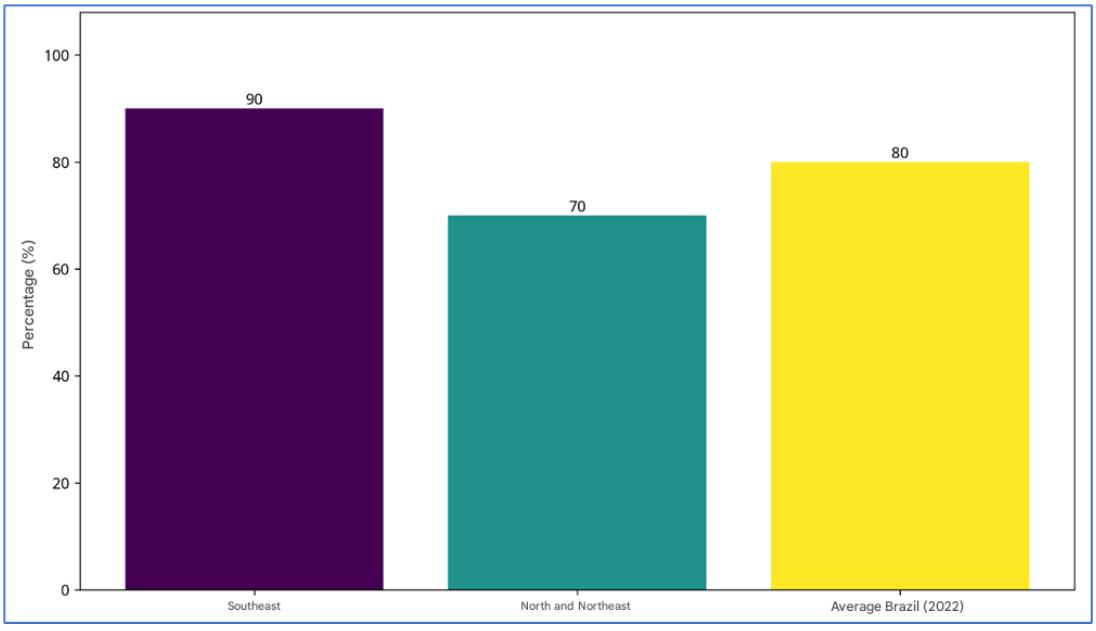
3.4.4.3 Regional Digital Inequalities

Brazil has relevant disparities in internet access that directly impact the equitable implementation of AI solutions. According to data from (IBGE, 2023)while almost 90% of households in the Southeast have access to the internet, the numbers drop to around 70% in the North and Northeast (Figure 4)(ESGINSIDE, 2025). The CGI.br's ICT Households 2024 survey revealed that only 22% of individuals have "satisfactory connectivity conditions", with only 3% in classes D and E having adequate connection conditions (Figure 5)(CGI.br, 2025; Observatório da Imprensa, 2024)

This digital inequality reproduces the social one, with 29 million Brazilians classified as "non-users" of the internet, 17 million of whom are black and 16 million from the lower classes (CGI.BR, 2025; Observatório Da Imprensa, 2024). In addition, the lack of digital literacy is a significant challenge, with only 30% of the Brazilian population possessing basic digital skills and less than 20% reaching an intermediate level of proficiency, which increases vulnerability to misinformation and fraud (Figure 6) (ESGINSIDE, 2025).

Figure 4

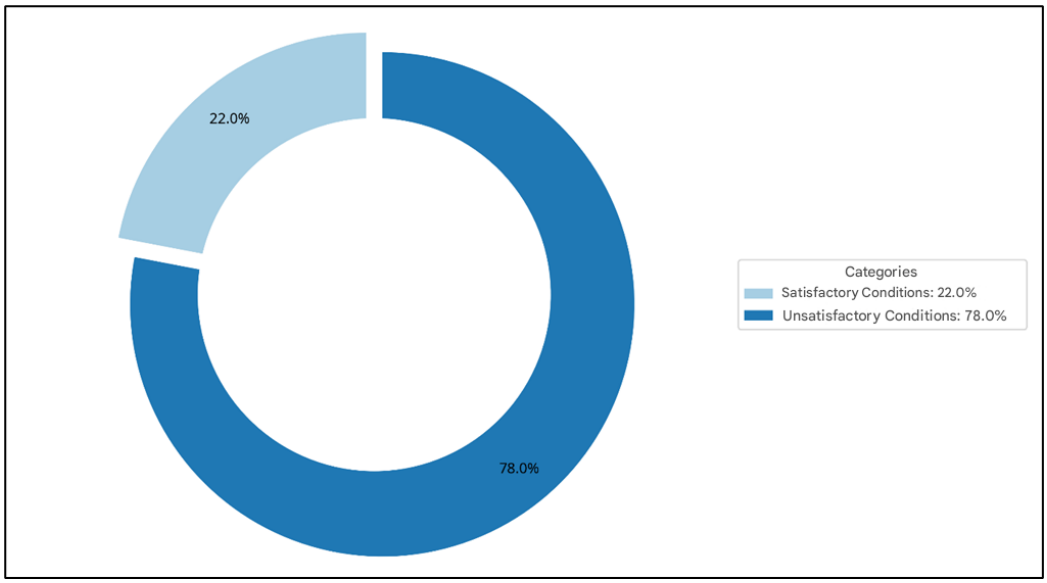
Internet Access in Households by Region (2022)



Source: (CGI.br, 2025; ESGINSIDE, 2025; IBGE, 2023).

Figure 5

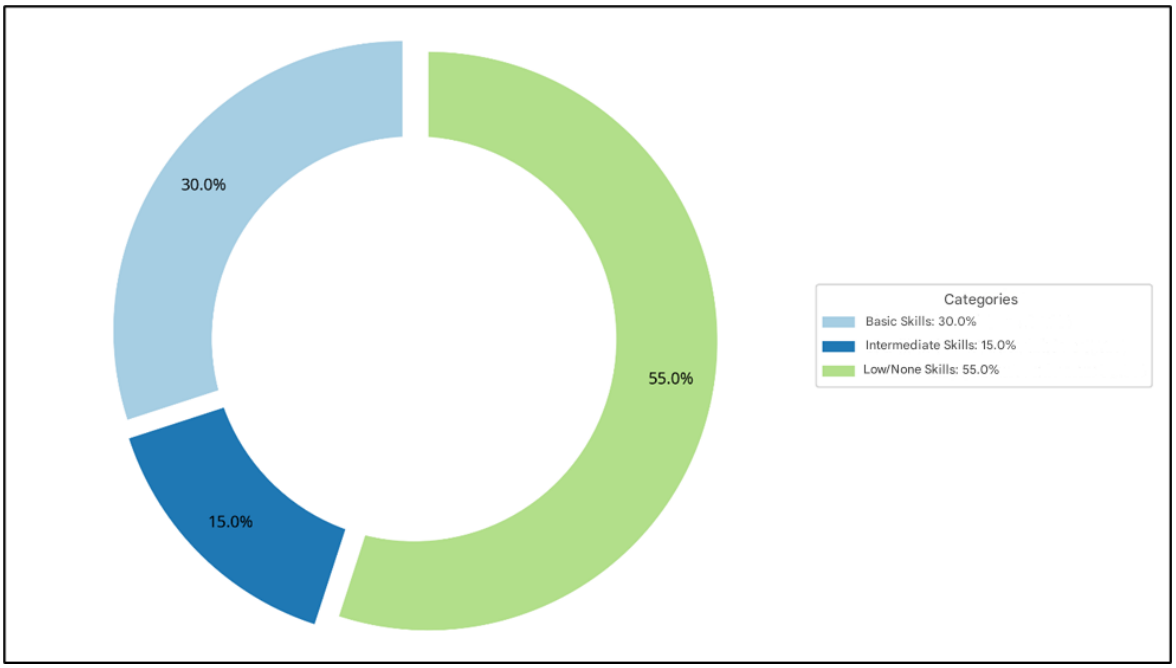
Satisfactory Connectivity Conditions in Brazil (2024)



Source: CGI.br(2025); Observatório da Imprensa(2024).

Figure 6

Digital Literacy Level in Brazil (2025)



Source: (ESGINSIDE, 2025)

Impactos Setoriais:

The regional digital divide has a different impact on the EBIA's priority economic sectors. In agribusiness, the implementation of Agro 4.0 faces connectivity limitations in rural areas, where internet access reached 84.8% of the population between 2016 and 2024 (Agência Gov, 2025). In the area of Health 4.0, regional disparities limit telemedicine and diagnostic AI systems, especially critical considering that 61.7% of the quilombola population lives in rural areas (IBGE, 2025).

For Smart Cities 4.0, the concentration of investments in fiber optics in the Southeast, versus dependence on unstable mobile networks in the North and Northeast, creates different capacities for the implementation of smart urban solutions (Esginside, 2025). In Industry 4.0, the lack of technological infrastructure and adequate connectivity prevents the full adoption of technologies such as IoT and AI, deepening social and regional inequalities in industrial modernization (Cardoso et al., 2017; Kubota; Rosa, 2023).

In the Education sector, the digital divide manifests itself in the lack of access to computers and stable internet, impacting the quality of education and widening the educational gap, especially in remote areas and among low-income populations (CNTE, 2023).

3.5 COMPARATIVE SYNTHESIS: GOVERNANCE FRAMEWORKS

In order to consolidate and structure the complex range of information discussed in the previous sections, the results regarding the governance of artificial intelligence are presented in a synthetic and visual way through three comparative tables. These tools were methodically developed to integrate and correlate the main findings of the central thematic axes investigated in this study: (1) to compare the regulatory approaches of the European Union, the United States and Brazil; (2) analyze and describe documented cases of algorithmic bias and its impacts on fundamental rights; (3) examine evidence on the effects of AI on the labor market, productivity, and inequalities; (4) contextualize the adoption of AI in Brazil and Latin America; and (5) propose recommendations for public policies for emerging countries that balance innovation, protection of rights, and mitigation of socioeconomic risks.

Table **1** articulates the main axes of social impact with specific and recommended governance instruments, connecting concrete problems to regulatory and technical solutions.

Table **2** offers a direct comparative analysis between the regulatory approaches of the European Union (with the AI Act), Brazil (with the LGPD and PL 2338/2023), and the United States (with NIST's AI RMF), highlighting their convergences and divergences. Table 3, in turn, establishes a timeline of the main global regulatory frameworks, illustrating the acceleration of debate and regulatory action. Together, these frameworks provide a cohesive and grounded picture that serves as a foundation for the proposed public policies for emerging countries, detailed in the subsequent section.

Table 1

Social impact pillars, case studies, and recommended governance tools

Impact Pillar	Central social risk	Documented cases	Recommended governance instruments	References
Algorithmic discrimination	Gender and racial bias	Amazon: Gender bias in recruitment, 2018; COMPAS: racial bias, accuracy 61%, black defendants 77% more likely to be classified as high risk; GPT-3.5: Gender bias in Portuguese-English translation	EIA mandatory for high risk; independent audits; data documentation/training; human supervision; bias testing; ConWea methodology for Portuguese	(Benatti et al., 2024; European Commission, 2025; OECD, 2024; ProPublica, 2025; Raulin & Angel, 2025; UNESCO, 2021)
Transparency and explainability	Opacity of models (black box)	Credit, recruitment, and criminal justice systems without explainability; GPAI models without transparency about capabilities/limitations	Transparency to users (chatbots, deepfakes); performance/limitation reports; XAI (Explainable AI); accessible technical documentation; Specific GPAI guidelines	(European Commission, 2025; European Parliament, 2019; NIST, 2023)
Privacy and data protection	Misuse of personal data	Generative AI trained on personal data without proper consent; Brazilian data in foreign clouds subject to the Cloud Act	LGPD compliance; legal bases; Minimizing; DPIA; third-party governance; anonymization; Government Cloud	(ANPD, 2024b, 2025a; Brasil, 2018; MGI, 2025a)

Work and reskilling	Occupational replacement / transformation	30% of jobs in the US could be automated by 2030; 60% will have tasks modified; women in more exposed administrative roles	Requalification programs; incentives for responsible adoption; sectoral partnerships; training vouchers; Women's Specific Tracks	(ILO, 2023; McKinsey Global Institute, 2023; National University, 2025; Stanford HAI, 2025)
General Purpose Models (GPAI)	Systemic externalities and emerging risks	Broad language models with unintended capabilities; risks of misinformation at scale; external technological dependence	Specific guides; safety tests; incident reporting; assessment of dangerous capabilities; red-teaming; Transparency Requirements for Foreign Models	(European Commission, 2025)
Adoption in the healthcare industry	Incorrect clinical decisions; Diagnostic bias	60% of Brazilian private hospitals use AI; Impact on 10 thousand to 140 thousand patients/month in laboratories	EIA for high-risk medical systems; certification; qualified human supervision; traceability of decisions; Specific clinical protocols	(European Parliament, 2019; PwC Brasil, 2025)
Electoral applications	Democratic manipulation; disinformation	Pioneering regulation of the TSE (Resolution 23,732/2024); Electoral deepfakes; chatbots simulating candidates	Mandatory labeling of AI content; prohibition of deepfakes; transparency in campaigns; audits of electoral systems	(BRASIL, 2024b; DFRLab & NetLab UFRJ, 2024)
Regional inequalities	Digital divide; Concentration of benefits	Southeast: 83% of households with internet vs. North: 81% (2024); South: 90%; Northeast: 81%; 61.7% of the quilombola population lives in rural areas (vs. 12.6% of the general population)	Priority digital inclusion; rural connectivity; digital training for traditional communities; Territorial impact assessment	(CGI.br, 2025; ESGINSIDE, 2025; Instituto Brasileiro de Geografia e Estatística (IBGE), 2025)
Technological sovereignty	External dependence; Data vulnerability	Brazilian data in foreign big techs; subjection to the American Cloud Act;	Government Cloud; development of national models; public-private	(Agência Brasil, 2025; MGI, 2025a; Penso, 2025; Silva et al., 2024)

		Concentration on foreign models	partnerships; South-South international cooperation	

Source: Compiled by the author based on Agência Brasil, 2025; ANPD, 2024B; ANPD, 2025; Brasil, 2018; Brasil, 2024A; CGI.BR, 2025; European Commission, 2025; DFRLAB NETLAB UFRJ, 2024; ESGINSIDE, 2025; IBGE, 2025; Mckinsey Global Institute, 2023; MGI, 2025; National University, 2025; NIST, 2023; OCDE, 2024; ILO, 2023; European Parliament, 2019; Penso Consultoria, 2025; Propublica, 2025; Pwc Brasil,2025; Silva et al., 2024; Stanford Hai, 2025; UNESCO, 2021.

Table 2

Synthetic comparison of regulatory frameworks (EU x Brazil x USA)

Dimension	European Union (AI Act)	Brazil (LGPD + PL 2338/2023 + Sectoral Initiatives)	United States (NIST AI RMF)
Approach	Risk-based prescriptive regulation	Hybrid: LGPD in force + risk-based bill in progress + sectoral regulations (TSE, Federal Revenue, CARF)	Voluntary framework + sectoral regulation
Risk Rating	4 levels: unacceptable, high, limited, minimum	Under development in PL 2338/2023; TSE set precedent for electoral AI	Context- and impact-based
Prohibited Uses	Social scoring, manipulation, vulnerability exploitation, real-time biometric recognition	*TSE: deepfakes, chatbots simulating candidates; Bill will define other prohibited uses*	There are no blanket prohibitions
High-risk systems	Strict obligations: EIA, data governance, documentation, transparency, human oversight	LGPD art. 20: review of automated decisions; PL proposes similar obligations; Federal Revenue: audit of tax algorithms	Risk management recommendations
GPAI Models	Specific guidelines (2025): transparency, testing, incident reporting	Under discussion; Dependence on foreign models being debated	No specific regulation
Implemented case studies	Gradual implementation in critical sectors	TSE (2024 elections), Federal Revenue Service (fraud detection), CARF (IARA system), hospitals (IA diagnosis)	Voluntary adoption by sectoral agencies

Dimension	European Union (AI Act)	Brazil (LGPD + PL 2338/2023 + Sectoral Initiatives)	United States (NIST AI RMF)
Enforcement	Fines of up to €35 million or 7% of global turnover	LGPD: fines up to 2% of revenue (limit BRL 50 million); PL proposes additional sanctions; TSE: electoral sanctions	No general sanctions; depends on sectoral regulation
Supervision	National authorities + European AI Office	ANPD (personal data) + AI authority to be defined; TSE (electoral); Serpro/Revenue (tax)	Sectoral agencies
Social participation	Public consultations on sectoral regulations	PL 2338/2023: more than 12 thousand contributions; OBIA inclusive seminars; Social Observatories	Consultations by specific agencies
Technological sovereignty	Focus on European companies and own models	Government Cloud; internal development (Serpro, Revenue); EBIA Applied Research Centers	Dependence on *national* big techs
Traditional communities	General protection of fundamental rights	AIIA Project (Indigenous appropriation of AI); 1.33 million quilombolas; Regional inequalities considered	Not specifically addressed

Source: Compiled by the author based on (Brasil, 2018; Câmara dos Deputados, 2025; European Commission, 2025; European Parliament, 2025; NIST, 2023; UNESCO, 2021; OCDE, 2024.
Note: GPAI = AI general-purpose models; *NIST AI RMF = NIST Artificial Intelligence Risk Management Framework.*

Table 3

Timeline of AI regulatory frameworks (2018–2025)

Year	Jurisdiction / Body	Instrument	Scope (law /standard/guide/ treaty/framework)	Key Notes	References
2018	Brazil	Law No. 13,709/2018 – LGPD	Law	Data protection base; material validity in 2020 and sanctions in 2021.	(Brasil, 2018)
2019	OCDE (Global)	Principles of AI	Guidelines/soft law	First widely endorsed international benchmark for trustworthy AI.	(OECD, 2024)
2020	European Union	White Paper on AI	White paper / consultation	It paves the way for the AI Act (risk approach).	(European Commission, 2020)
2021	Brazil	Brazilian AI Strategy (EBIA)	National guidelines	Guidelines for innovation and responsible use.	(Brasil, 2021)
2021	UNESCO (Global)	Recommendation on AI Ethics	Multilateral recommendation	Human rights, dignity, justice, human oversight.	(UNESCO, 2022)
2022	USA	Blueprint for an AI Bill of Rights	Rights Guide (Non-Binding)	Safeguards for automated systems targeting people.	(United States, 2022)
2022	European Union	AI Liability Directive (proposal)	Directive (procedure)	Civil liability rules complementary to the AI Act.	(European Parliamentary Research Service, 2023)
2022	China	Algorithmic Recommendation Services Rules	Sectoral regulation	Algorithmic recommendation transparency and governance.	(China. Central Cyberspace Affairs Commission Office, 2021)
2023	ISO/IEC (Global)	ISO/IEC 23894:2023 – Risk Management for AI	International Standard	AI risk management in organizations.	(ISO & IEC, 2023)
2023	USA (NIST)	AI RMF 1.0	Volunteer Framework	Identification, assessment and mitigation of risks throughout the life cycle.	(NIST, 2023)
2023	USA	Executive Order 14110 (Secure and Reliable AI)	Executive Order	Testing of security, reporting, governance in GPAI and use in the public sector.	(United States, 2023)
2023	G7 (Global)	Hiroshima AI Process – Code of Conduct	Code of Conduct	Guidance for advanced model developers.	(European Commission, 2023)
2023	United Kingdom (Global)	Bletchley Declaration	Multilateral declaration	Frontier AI Risk Recognition.	(UK government, 2025))

2023	ISO/IEC (Global)	ISO/IEC 42001:2023 – AI Management System (AIMS)	Management System Standard	AI management requirements (analogous to ISO 9001 for AI).	(ISO, 2023)
2024	European Union	AI Act	Regulation	First horizontal regime binding by risk levels; includes GPAI.	(Future of Life Institute, 2025)
2024	Council of Europe	Treaty on AI (CAI)	International treaty	Human rights applied to AI; open to non-European countries.	(Council of Europe, 2024)
2024	Brazil (TSE)	Resolution No. 23,732/2024	Electoral resolution	Rules on AI/deepfakes, labeling, and accountability in campaigns.	(Brasil, 2024a)
2024	Brazil (ANPD)	Technical Note 27/2024	Technical Guidance	Automated decisions, generative AI, regulatory agenda.	(ANPD, 2024a)
2024	Brazil (ANPD)	Technical Note 30/2024	Technical Guidance	Automated decisions, generative AI, regulatory agenda.	(ANPD, 2024b)
2024	OCDE (Global)	AI Principles – Update	Guidelines/soft law	Incorporates advanced model risks and strengthened governance.	(OECD, 2024)
2025	Brazil (ANPD)	Technical Note 12/2025	Technical Guidance	Automated decisions, generative AI, regulatory agenda.	(ANPD, 2025a)
2025	Brazil (ANPD)	Technological radar	Technical Guidance	Automated decisions, generative AI, regulatory agenda.	(ANPD, 2025b)
2025	Brazil	PL 2338/2023 (expected approval)	Bill	Legal framework for AI in the country with a focus on ethical and responsible use.	(Câmara dos Deputados, 2025)
2025	USA	Removing barriers to american leadership in artificial intelligence	Executive Order	*Repeals Executive Order 14110 and prioritizes U.S. global leadership in AI through free markets, without "ideological biases" or "social agendas."*	(United States, 2025)

Source: Compiled by the author based on ANPD, 2024a; ANPD, 2024b; ANDP, 2025a; ANDP, 2025b; Brasil, 2024a; Câmara Dos Deputados, 2025; United States, 2025.

3.6 PROPOSAL OF PUBLIC POLICIES FOR EMERGING COUNTRIES

The comparative analysis of regulatory frameworks (Sections 3.1 and 3.2), documented cases of algorithmic bias (Section 3.3), and evidence on socioeconomic impacts (Section 3.4) allows us to propose an integrated framework of public policies for emerging countries. The proposal presented in

Table **4** is structured around five complementary pillars that respond directly to the identified gaps: risk-based regulation (inspired by the AI Act and PL 2338/2023), algorithmic impact assessment (based on the Amazon and COMPAS cases), institutional strengthening (meeting capacity limitations), work and reskilling policies (responding to automation projections), and transparency with social participation (incorporating multilateral recommendations).

Table 4

Proposal of Public Policies for the use of AI in Emerging Countries

Pillar	Políticas Publicas	Key References
1. Context-Adapted Risk-Based Regulation	• Adoption of a risk-based approach similar to the AI Act, but with requirements commensurate with institutional capacity. • Prioritization of critical sectors: health, education, justice, credit and employment. • Clear definition of prohibited uses in line with constitutional values. • Regulatory *sandbox* mechanisms for controlled innovation.	(Câmara dos Deputados, 2025; European Commission, 2025; ProPublica, 2025; Tse et al., 2024)
2. Mandatory Algorithmic Impact Assessment (EIA)	• EIA required for high-risk systems before implementation. • Standardized methodology including: bias analysis, impact on fundamental rights, mitigation measures and monitoring plans. • Independent Audits Proportional to Risk • Publication of executive summaries for public transparency.	(Câmara dos Deputados, 2025; European Commission, 2025; ProPublica, 2025; Tse et al., 2024)
3. Institutional Strengthening and Capacity Building	• Investment in technical training of regulatory authorities. • Creation of centers of excellence in AI governance. • Partnerships with universities for applied research. • International cooperation to share best practices	(NIC.br, 2025; NUPEPA, 2025; Pires et al., 2023; Silva et al., 2024)
4. Active Work and Requalification Policies	• Mapping of exposed occupations and development of requalification trails. • Training vouchers for workers in high-exposure sectors.	(CNTE, 2023; McKinsey Global Institute, 2023; Santos, 2023)

	• Tax incentives for companies that invest in requalification. • Specific programs for women in administrative roles. • Strengthening STEM education and digital literacy.	
5. Transparency and Social Participation	• Transparency requirements for GPAI models used locally. • Mechanisms for public participation in regulatory decisions. • Algorithmic education in school curricula. • Accessible channels for reporting algorithmic discrimination. • Independent observatories for impact monitoring.	(Câmara dos Deputados, 2025; Data Privacy Brasil, 2025; NIST, 2023; Pretto, 2024)

Source: Compiled by the author based on Câmara dos Deputados, 2025; European Commission, 2025; CNTE, 2023; Data Privacy Brasil, 2025; NIC.BR, 2025; NIST, 2023; NUPEPA, 2025; Mckinsey Global Institute, 2023; Pires et al., 2023; Pretto, 2024; Propublica, 2025; Santos, 2023; Silva et al., 2024; TSE et al., 2024.
Note: The proposed policies consider the specificities of emerging countries, including limited institutional capacity, regional inequalities, and the need to balance innovation with the protection of fundamental rights

4 DISCUSSION

The results presented show global convergence around fundamental ethical principles for AI governance, transparency, equity, accountability, human oversight and protection of rights, but significant divergence in the regulatory approaches adopted by different jurisdictions. This section discusses the implications of these findings, analyzes tensions and trade-offs inherent in AI governance, and proposes pathways for emerging countries.

4.1 ALGORITHMIC BIAS AS A SYSTEMIC CHALLENGE

The Amazon and COMPAS cases illustrate that algorithmic bias is not an isolated technical failure, but a systemic problem that reflects and amplifies existing social inequalities. As he notes (Chen, 2023, p. 3–4) algorithmic bias stems from both limitations in training datasets and biased decisions by developers. When AI systems are trained with historical data that reflects past discrimination, such as mostly male resumes in the tech industry or racially biased court rulings, they learn and perpetuate these patterns (Liefgreen et al., 2024)

ProPublica's finding that COMPAS maintained racial bias even after controlling for confounding variables is particularly troubling, as it suggests that the problem cannot be solved by simply "balancing" *datasets* or removing sensitive variables. Proxy variables, such as zip code, school type, or language patterns, can serve as substitutes for protected traits, allowing bias to persist in more subtle and hard-to-detect ways (Angwin et al., 2016)

In the Brazilian context, where structural inequalities of race, gender, and class are deep and historically rooted, the risk of AI systems amplifying these disparities is substantial (Chen, 2023) The high rate of enterprise adoption (40% of companies) without consolidated regulatory frameworks creates a scenario of vulnerability (AWS & Strand Partners, 2025) especially when we consider that many systems are developed and trained with data from foreign contexts that may not adequately reflect the Brazilian reality, in a country where only 32% of companies consider themselves prepared with the current set of competencies to fully leverage AI (AWS & Strand Partners, 2025)

4.2 TENSIONS BETWEEN INNOVATION AND REGULATION

The comparison between the European, American and Brazilian approaches reveals different philosophies about the balance between promoting innovation and protecting rights. The European AI Act adopts a stance of caution, establishing preventive obligations for high-risk systems before their implementation. This approach offers greater predictability and protection, but it can create barriers to entry for startups and small businesses, potentially concentrating the market on large players with resources for compliance.

The American model, based on voluntary frameworks and sectoral regulation, offers greater flexibility and can accelerate innovation, but relies heavily on self-regulation and can leave protection gaps, especially for vulnerable populations. The absence of blanket bans for high-risk uses contrasts with the European approach and may allow for problematic applications.

Brazil is in an intermediate position, with the LGPD providing a solid basis for personal data protection and PL 2338/2023 proposing a risk-based approach similar to the European one, but adapted to the national context and institutional realities. This position offers a unique opportunity to learn from international experiences, avoiding both regulatory overreach and protection gaps through calibrated, evidence-based implementation.

4.3 SPECIFIC CHALLENGES OF EMERGING COUNTRIES

Emerging countries like Brazil face distinct challenges in AI governance:

- Limited institutional capacity: Effective implementation of complex regulations requires significant human, technical, and financial resources. The ANPD, for example, operates with a limited budget and staff compared to European authorities, making it difficult to oversee comprehensive;

- Technological dependency: Most large-scale AI models are developed by foreign companies, creating power asymmetries and making it difficult to enforce national regulations. Training data may not adequately reflect local contexts, increasing risks of bias;
- Investment gap: Although Brazil leads Latin America in adoption, it invests only 30% of its economic potential in AI. This gap limits the ability to develop domestic solutions and increases dependence on foreign technologies;
- Structural inequalities: Deep inequalities in access to education, technology, and economic opportunity mean that the benefits of AI can be concentrated in elites, while risks (such as job automation) disproportionately affect vulnerable populations.

4.4 GENERAL PURPOSE MODELS: EMERGING CHALLENGE

General purpose models (GPAI) represent a unique regulatory challenge. Unlike specific AI systems, GPAIs are developed for multiple applications, many of which developers did not envision. This versatility makes it difficult to assess preventive risk and assign responsibilities (Hendrycks et al., 2023; Syuntyurenko, 2022) Consequently, the lack of a clear and comprehensive definition for GPAIs in the regulatory context can result in significant gaps in protection and accountability, which can undermine regulatory effectiveness and user safety.

The 2025 European guidelines for GPAI recognize this complexity, explicitly establishing transparency obligations on capabilities and limitations, rigorous safety testing procedures and timely reporting of major incidents to relevant regulatory supervisory competent authorities.(European Commission, 2025). For emerging countries, the question is how to regulate GPAIs developed abroad but widely used locally. Possible approaches include transparency requirements for providers, impact assessments for high-risk local applications, and international cooperation for sharing information on incidents (Kiseleva et al., 2022; Kuteynikov et al., 2020) These guidelines aim to ensure that the implementation of GPAIs not only promotes innovation but also protects the fundamental rights of users.

4.5 WORK AND RESKILLING: URGENCY OF ACTIVE POLICIES

Evidence on labour market impacts indicates an urgent need for active employment and reskilling policies. The ILO's identification of greater exposure of tasks performed by

women suggests that, without intervention, AI can exacerbate existing gender inequalities in the labour market (ILO, 2023).

Effective reskilling programs should be: (1) Targeted: focused on occupations and groups that are more exposed; (2) Affordable: with vouchers, scholarships and flexibility for employed workers; (3) Prospective: anticipating future demands, not just reacting to displacements; (4) Integrated: combining technical skills with socio-emotional competencies valued in AI contexts.

Brazil's scenario of high adoption but underinvestment suggests that many companies are implementing AI without adequate employee preparation, creating risks of exclusion and resistance. In addition, the lack of a clear strategy for AI integration can result in significant challenges in the acceptance and effectiveness of implemented solutions (Campos Zabala, 2023)

4.6 ETHICAL DIMENSION AND SOCIAL PARTICIPATION IN AI GOVERNANCE

4.6.1 Algorithmic Bias in Language Models in Portuguese

Specific Challenges of Brazilian Portuguese:

Recent research has identified significant gender bias in machine translation models into Portuguese. (Raulin & Angel, 2025) evaluated GPT-3.5 turbo using the Challenge Test Set WinoMT translated into Brazilian Portuguese, demonstrating that the model tends to promote gender bias in the translation of nouns related to professions. The adaptation of the ConWea approach to classifying texts in Portuguese revealed the perpetuation of gender stereotypes in scientists' biographies and social media content (BENATTI et al., 2024). These findings indicate that the biases present in datasets predominantly in English are reproduced and amplified when applied to the Brazilian linguistic and cultural context.

Mitigation Mechanisms:

The automatic detection of bias in the Brazilian corpus has been developed through natural language processing techniques (Raulin & Angel, 2025) The national literature points to the need to incorporate AI literacy into school curricula and raise awareness to improve the capacity for critical evaluation of these technologies (Santos, 2023) Recent studies on language models in Portuguese emphasize the importance of training with diverse and representative datasets of Brazilian Portuguese to ensure more accurate and unbiased results (Corrêa, 2024; Pires et al., 2023)

4.6.2 Perspectives of Traditional Communities and Vulnerable Populations

Technological Appropriation Initiatives:

The project "AIIA – Indigenous appropriation of Artificial Intelligence" represents a pioneering example of how traditional communities can critically appropriate AI (IberCultura Viva, 2024) Nine indigenous people from different peoples in Brazil, Argentina and Chile collaborate in the production of 10 works of contemporary indigenous art using tools such as Midjourney and Adobe Firefly. This initiative demonstrates that the conscious, critical, and technical appropriation of AI by traditional communities can serve as a mechanism for cultural preservation and projection, countering narratives of technological substitution.

Differentiated Impacts:

The 2022 Census counted 1,330,186 quilombola people in Brazil, with 61.7% living in rural areas compared to 12.6% of the national population (IBGE, 2025) This geographical distribution implies differentiated access to AI technologies, with potential impacts on the provision of digital public services. The literature on traditional peoples and biodiversity indicates that indigenous and quilombolas possess ancestral knowledge that can be complementary to AI approaches to sustainability, particularly in areas such as pollinator protection and environmental conservation (SBPC, 2024)

4.6.3 Social Participation Mechanisms in AI Governance

Experience of PL 2338/2023:

Bill 2338/2023, which establishes the legal framework for artificial intelligence in Brazil, went through a public consultation process that received more than 12 thousand contributions from civil society (Câmara dos Deputados, 2025). This process demonstrates Brazilian institutional capacity to incorporate social participation in AI policymaking, setting a precedent for future participatory governance mechanisms.

Applicable International Models:

The international literature points to three concrete measures to strengthen civil society involvement in AI governance: (1) more open and transparent selection processes for participation in expert groups; (2) imposition of the duty to respond to civil society contributions with clear feedback mechanisms; and (3) adequate funding to reduce barriers to participation (Data Privacy Brasil, 2025). The Brazilian experience with the Brazilian Public Software Portal, which completed 20 years, offers a consolidated model of social participation in technological policies that can be adapted to AI (Pretto, 2024).

AI Social Observatories:

The Brazilian Observatory of Artificial Intelligence (BOAI), linked to the Brazilian Plan for Artificial Intelligence, represents a structured initiative for participatory monitoring (NIC.BR, 2025). OBIA conducts seminars with national and international experts on the impacts of AI on the labor market, education, and sustainable development, promoting inclusive and reasoned dialogue. Additionally, initiatives such as the Social Observatory for Artificial Intelligence and Digital Data (OSIADD) produce reports on perceptions, uses, and impacts of AI in the Social Sciences and Humanities, connecting Brazilian and Portuguese researchers (NUPEPA, 2025)

4.6.4 Data Sovereignty and Technology Dependency

National Context:

The issue of digital sovereignty emerges as a strategic imperative in the face of dependence on foreign technological infrastructures. Brazilian data, including information from the IBGE and National Security, are stored in big techs and non-national clouds, creating vulnerabilities to attacks and external pressures (Queiroz, 2024) The American Cloud Act allows the US government to demand access to data from American companies, even when physically stored in Brazil, conflicting with LGPD principles (Penso, 2025)

Sovereignty Initiatives:

The Ministry of Management and Innovation consolidated the Government Cloud as a milestone of technological sovereignty, in partnership with Serpro and Dataprev (MGI, 2025b). This initiative ensures that data from public services and policies are processed in infrastructure under state management, benefiting strategic areas such as Health, Education and Social Security. The Network for Digital Sovereignty prepared a public letter to President Lula reinforcing the importance of a Digital Plan for National Sovereignty, highlighting the need to build national clouds and regulate platforms (Rede pela Soberania Digital, 2024).

5 CONCLUSIONS

This integrative review showed that effective governance of artificial intelligence requires a multidimensional approach that integrates risk-based regulation, operational management frameworks, ethical principles based on human rights, and active socioeconomic policies. Global convergence around principles such as transparency, equity,

accountability, and human oversight provides a solid basis for coordinated action, but the diversity of regulatory approaches reflects different institutional priorities and capacities.

The documented cases of algorithmic bias, Amazon and COMPAS, demonstrate that AI systems can perpetuate and amplify existing inequalities when developed without proper governance. In the Amazon case, gender bias emerged from biased historical data and manifested itself in explicit penalties for female candidates. In the COMPAS case, racial bias persisted even after controlling for confounding variables, resulting in black defendants being misclassified as high-risk at rates twice as high as white defendants. These examples reinforce the need for mandatory algorithmic impact assessments, independent audits and qualified human oversight, especially in applications that affect fundamental rights.

The Brazilian context has unique characteristics that should inform public policies. Brazil leads Latin America in AI adoption, with 40% of companies already using the technology systematically, but investing only 30% of its economic potential. This combination of high adoption and underinvestment, coupled with the absence of consolidated regulatory frameworks, creates significant vulnerabilities. The LGPD provides a solid basis for personal data protection, and PL 2338/2023 proposes a risk-based approach similar to the European one, but adapted to the national context. The effective implementation of these regulations will require institutional strengthening, investment in technical training, and social participation mechanisms.

The impacts on the labor market require urgent attention. Evidence indicates that up to 30% of current jobs in the U.S. could be automated by 2030, and 60% will have significantly modified tasks. The ILO identifies greater exposure of women in administrative roles, suggesting that without intervention, AI can exacerbate gender inequalities. Active reskilling policies, targeting occupations and groups that are most exposed, are essential to ensure a just and inclusive transition.

For emerging countries, public policy recommendations are proposed structured on five pillars: (1) risk-based regulation adapted to the context; (2) mandatory algorithmic impact assessment for high-risk systems; (3) institutional strengthening and technical training; (4) active work and requalification policies; and (5) transparency and social participation. This framework seeks to balance the promotion of innovation with the protection of rights, recognizing limitations of institutional capacity and the need for strategic prioritization.

The governance of general purpose models (GPAI) represents an emerging challenge that requires international cooperation. The 2025 European guidelines set an important

precedent, but emerging countries should develop their own mechanisms to regulate GPAIs developed abroad but widely used locally, including transparency requirements, impact assessments for high-risk applications, and sharing information about incidents.

In summary, evidence-based governance, anchored in fundamental rights (LGPD), multilateral principles (UNESCO/OECD), risk management (NIST) and risk-based model (AI Act), combined with algorithmic impact assessments, proportional audits, GPAI transparency and reskilling policies, constitutes a promising path to ensure that AI promotes equity, accountability and innovation with a public purpose. For Brazil and other emerging countries, the moment is one of opportunity: to learn from international experiences, adapt solutions to local contexts, and build AI ecosystems that balance economic development, social inclusion, and respect for fundamental rights.

Future research should deepen: (1) empirical evaluation of the effectiveness of different regulatory instruments; (2) development of algorithmic impact assessment methodologies adapted to resource-constrained contexts; (3) analysis of the distributive impacts of AI in emerging countries, with attention to race, gender, and class intersectionalities; (4) study of international cooperation mechanisms for GPAI governance; and (5) evaluation of reskilling programs and their success rates in facilitating occupational transitions.

REFERENCES

ABES. (2025). *Mercado Brasileiro de Software – Panorama e Tendências 2025*. Associação Brasileira das Empresas de Software. https://abes.org.br/dados-do-setor/

Agência Brasil. (2025, julho). *Soberania digital do Brasil requer autonomia tecnológica, diz deputado*. https://agenciabrasil.ebc.com.br/politica/noticia/2025-07/soberania-digital-do-brasil-requer-autonomia-tecnologica-diz-deputado

Agência Gov. (2025). Acesso à internet nas áreas rurais dispara 150% em oito anos e alcança 84,8% da população. *Agência Gov*. https://agenciagov.ebc.com.br/noticias/202507/acesso-a-internet-nas-areas-rurais-dispara-150-em-oito-anos-e-alcanca-84-8-da-populacao

Amazon Web Services. (2025). *Pesquisa sobre adoção de IA na América Latina*. AWS. https://www.aboutamazon.com.br/noticias/noticias-da-empresa/pesquisa-da-aws-mostra-que-brasil-lidera-adocao-de-ia-na-america-latina

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine Bias: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*.

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

ANPD. (2024a). *Nota Técnica nº 27/2024/FIS/CGF/ANPD – Tratamento de dados pessoais de terceiros com finalidade de desenvolver modelo de inteligência artificial generativa*. Autoridade Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/centrais-de-conteudo/documentos-tecnicos-orientativos/nt-27.pdf

ANPD. (2024b). *Nota Técnica nº 30/2024/CGN/ANPD – Proposta de Agenda Regulatória para o biênio 2025–2026*. Autoridade Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/acesso-a-informacao/participacao-social/outras-acoes/documentos/nota-tecnica-no-30-2024-cgn-anpd.pdf

ANPD. (2025a). *Nota Técnica nº 12/2025/CON1/CGN/ANPD – Consolidação das contribuições recebidas na Tomada de Subsídios – Inteligência Artificial e Revisão de Decisões Automatizadas*. Autoridade Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/acesso-a-informacao/participacao-social/outras-acoes/documentos/ts-06-2024-nt-12-2025-consolidacao-das-contribuicoes.pdf/@@download/file

ANPD. (2025b, outubro). *ANPD publica 5ª Edição do Radar Tecnológico*. Agência Nacional de Proteção de Dados. https://www.gov.br/anpd/pt-br/assuntos/noticias/anpd-publica-5a-edicao-do-radar-tecnologico

AWS & Strand Partners. (2025). *Desbloqueando o Potencial da IA no Brasil*. Amazon Web Services. https://www.aboutamazon.com.br/noticias/noticias-da-empresa/pesquisa-da-aws-mostra-que-brasil-lidera-adocao-de-ia-na-america-latina

Benatti, R., Severi, F., Avila, S., & Colombini, E. L. (2024). Gender Bias Detection in Court Decisions: A Brazilian Case Study. *The 2024 ACM Conference on Fairness Accountability and Transparency*, 746–763. https://doi.org/10.1145/3630106.3658937

Brasil. (2018). *Lei nº 13.709, de 14 de agosto de 2018 – Lei Geral de Proteção de Dados Pessoais (LGPD)*. Planalto. https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm

Brasil. (2021). *Portaria nº 4.617, de 6 de abril de 2021*. Ministério da Ciência, Tecnologia e Inovações. https://antigo.mctic.gov.br/mctic/opencms/legislacao/portarias/Portaria_MCTI_n_4617_de_06042021.html

Brasil. (2024a). *Receita desenvolve ferramenta inovadora capaz de ampliar detecção de fraudes tributárias e aduaneiras*. Ministério da Fazenda. https://www.gov.br/fazenda/pt-br/assuntos/noticias/2024/setembro/receita-desenvolve-ferramenta-inovadora-capaz-de-ampliar-deteccao-de-fraudes-tributarias-e-aduaneiras

BRASIL. (2024b). *Resolução nº 23.732, de 27 de fevereiro de 2024*. Tribunal Superior Eleitoral. https://www.tse.jus.br/legislacao/compilada/res/2024/resolucao-no-23-732-de-27-de-fevereiro-de-2024

Brasil. (2025). *CARF e FGV assinam contrato de encomenda tecnológica para aprimoramento de IA em julgamentos administrativos fiscais*. Conselho Administrativo de Recursos Fiscais. https://www.gov.br/carf/pt-br/assuntos/noticias/carf-e-fgv-assinam-contrato-de-encomenda-tecnologica-para-aprimoramento-de-ia-em-julgamentos-administrativos-fiscais

Brookings Institution. (2025). *New data show no AI jobs apocalypse—For now*. Brookings. https://www.brookings.edu/articles/new-data-show-no-ai-jobs-apocalypse-for-now/

Câmara dos Deputados. (2025). *Projeto de Lei nº 2338/2023 – Marco Legal da IA no Brasil (tramitação)*. Câmara dos Deputados. https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2487262

Campos Zabala, F. J. (2023). Integrating AI into Existing Systems and Processes. Em F. J. Campos Zabala, *Grow Your Business with AI* (p. 413–431). Apress. https://doi.org/10.1007/978-1-4842-9669-1_18

Cardoso, W., Azzolini Junior, W., Bertosse, J. F., Bassi, E., & Ponciano, E. S. (2017). Digital manufacturing, industry 4.0, clould computing and thing internet: Brazilian contextualization and reality. *Independent Journal of Management & Production*, *8*(2), 459. https://doi.org/10.14807/ijmp.v8i2.572

CGI.br. (2025). *Pesquisa TIC Domicílios 2024*. Comitê Gestor da Internet no Brasil. https://cetic.br/media/docs/publicacoes/2/20250512120132/tic_domicilios_2024_livro_eletronico.pdf

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, *10*(1), 1–16. https://doi.org/10.1057/s41599-023-02079-x

China. Central Cyberspace Affairs Commission Office. (2021). *Regulation on the Administration of Algorithm Recommendations for Internet Information Services*. Cyberspace Administration of China, Ministry of Industry and Information Technology, Ministry of Public Security, State Administration for Market Regulation. https://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm

CNTE. (2023). Desigualdade no acesso à internet impacta qualidade da educação. *CNTE Notícias*. https://cnte.org.br/noticias/desigualdade-no-acesso-a-internet-impacta-qualidade-da-educacao-c1bf

Corrêa, N. K. (2024). Open-source tiny language models trained in Brazilian Portuguese. *Software Impacts*, *20*. https://doi.org/10.1016/j.simpa.2024.100343

Costa, A. F. J., Cebrian, F. F., dos Anjos, L. C., Guedes, M. S., & Moraes, T. G. (2024). *Radar Tecnológico nº 3: Inteligência Artificial Generativa* (Publicação digital – PDF No. 3; Radar Tecnológico). Autoridade Nacional de Proteção de Dados (ANPD). https://www.gov.br/anpd

Council of Europe. (2024, setembro). *Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*. https://rm.coe.int/1680afae3c

Dastin, J. (2018). Insight—Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/

Data Privacy Brasil. (2025, fevereiro). *Governança global de IA: agindo com a sociedade civil*. https://www.dataprivacybr.org/governanca-global-de-ia-agindo-com-a-sociedade-civil/

DFRLab & NetLab UFRJ. (2024, maio). *A regulação do uso de IA para as eleições brasileiras. O que está em jogo?* NetLab UFRJ. https://netlab.eco.ufrj.br/post/a-regula%C3%A7%C3%A3o-do-uso-de-ia-para-as-elei%C3%A7%C3%B5es-brasileiras-o-que-est%C3%A1-em-jogo

ECLAC. (2025, outubro). *Latin America and the Caribbean accelerate the adoption of artificial intelligence, though challenges remain in investment, talent, and governance*. Economic Commission for Latin America and the Caribbean. https://www.cepal.org/en/pressreleases/latin-america-and-caribbean-accelerate-adoption-artificial-intelligence-though

Economist Impact & JPMorgan. (2024, agosto). *The great leap: Harnessing genAI to revolutionize Latin America's service economy*. J.P. Morgan Private Bank. https://privatebank.jpmorgan.com/latam/en/insights/markets-and-investing/the-great-leap-harnessing-gen-ai-to-revolutionize-latin-americas-service-economy

ESGINSIDE. (2025). Desigualdade digital e o desafio da conectividade no Brasil. *ESG Inside*. https://esginside.com.br/2025/05/12/desigualdade-digital-e-o-desafio-da-conectividade-no-brasil/

European Commission. (2020). *White Paper on Artificial Intelligence—A European approach to excellence and trust*. https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Commission. (2023, outubro). *Hiroshima Process International Code of Conduct for Advanced AI Systems*. https://digital-strategy.ec.europa.eu/en/library/hiroshima-process-international-code-conduct-advanced-ai-systems

European Commission. (2025). *AI Act – Regulatory framework for Artificial Intelligence (inclui diretrizes para modelos de propósito geral)*. European Commission. https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

European Parliament. (2019). *A governance framework for algorithmic accountability and transparency* (Study No. EPRS_STU(2019)624262). European Parliamentary Research Service. https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2019)624262

European Parliament. (2025). *EU AI Act: First regulation on artificial intelligence (timeline e materiais)*. Europian Parliament. https://www.europarl.europa.eu

European Parliamentary Research Service. (2023). *Artificial intelligence liability directive*. European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf

European Union. (2016, maio 4). *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Felten, E. W., Raj, M., & Seamans, R. (2023). *Occupational heterogeneity in exposure to generative AI*. https://doi.org/10.2139/ssrn.4414065

Future of Life Institute. (2025). *EU Artificial Intelligence Act | Up-to-date developments and analyses of the EU AI Act*. https://artificialintelligenceact.eu/

Hendrycks, D., Mazeika, M., & Woodside, T. (2023). *An Overview of Catastrophic AI Risks* (Versão 6). arXiv. https://doi.org/10.48550/ARXIV.2306.12001

IberCultura Viva. (2024). *AIIA – Apropriação Indígena da Inteligência Artificial*. https://iberculturaviva.org/cultura-quilombola-fe-e-folia-arte-indigena-e-inteligencia-artificial-os-projetos-do-brasil-selecionados-no-edital-de-apoio-a-redes-2023/

IBGE. (2023). *Pesquisa Nacional por Amostra de Domicílios Contínua—PNAD Contínua: Principais destaques da evolução do mercado de trabalho no Brasil 2012-2022*. Instituto Brasileiro de Geografia e Estatística (IBGE). https://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Principais_destaques_PNAD_continua/2012_2022/PNAD_continua_retrospectiva_2012_2022.pdf

IBGE. (2025). *Censo Demográfico 2022: Quilombolas—Primeiros resultados*. Instituto Brasileiro de Geografia e Estatística. https://www.ibge.gov.br/estatisticas/sociais/populacao/37225-censo-demografico-2022-quilombolas-primeiros-resultados.html

Ilia Digital. (2025). *Inteligência Artificial no Brasil: 9 Milhões de Empresas já Adotaram IA em 2025*. Ilia. https://ilia.digital/inteligencia-artificial-no-brasil-9-milhoes-de-empresas-ja-adotaram-ia-em-2025/

ILO. (2023). *Generative AI and Jobs: A global analysis of potential effects on the quantity and quality of jobs*. The International Labour Organization. https://www.ilo.org

Instituto Brasileiro de Geografia e Estatística (IBGE). (2025). Censo 2022: 61,7% da população quilombola vive em áreas rurais. *Agência de Notícias IBGE*. https://agenciadenoticias.ibge.gov.br/agencia-noticias/2012-agencia-de-

noticias/noticias/43313-censo-2022-94-6-dos-quilombolas-em-areas-rurais-convivem-com-precariedades-no-saneamento-basico

ISO. (2023). *ISO/IEC 42001:2023—Information technology—Artificial intelligence—Management system*. https://www.iso.org/standard/42001

ISO & IEC. (2023). *ISO/IEC 23894:2023—Information technology—Artificial intelligence—Guidance on risk management*. International Organization for Standardization. https://www.iso.org/standard/77304.html

Jung, J., & Katz, R. L. (2025). *Impacto económico de la inteligencia artificial en América Latina: Transformación tecnológica y rezago en materia de inversión y capacidades laborales* (Documentos de Proyectos No. LC/TS.2025/37/Rev.1; p. 47). Comissão Econômica para a América Latina e o Caribe. https://www.cepal.org/es/publicaciones/81909-impacto-economico-la-inteligencia-artificial-america-latina-transformacion

Kiseleva, A., Kotzinos, D., & De Hert, P. (2022). Transparency of AI in Healthcare as a Multilayered System of Accountabilities: Between Legal Requirements and Technical Limitations. *Frontiers in Artificial Intelligence*, *5*, 879603. https://doi.org/10.3389/frai.2022.879603

Kubota, L. C., & Rosa, M. B. (2023). TD 2867 - Internet das coisas no Brasil: Análise e recomendação de políticas com ênfase no agronegócio. *Texto para Discussão*, 1–45. https://doi.org/10.38116/td2867

Kuteynikov, D. L., Izhaev, O. A., Zenin, S. S., & Lebedev, V. A. (2020). Algorithmic Transparency and Accountability: Legal Approaches to Solving the “Black Box” Problem. *Lex Russica*, *73*(6), 139–148. https://doi.org/10.17803/1729-5920.2020.163.6.139-148

Liefgreen, A., Weinstein, N., Wachter, S., & Mittelstadt, B. (2024). Beyond ideals: Why the (medical) AI industry needs to motivate behavioural change in line with fairness and transparency values, and how it can do it. *AI & SOCIETY*, *39*(5), 2183–2199. https://doi.org/10.1007/s00146-023-01684-3

McKinsey Global Institute. (2023). *Generative AI and the future of work in America*. McKinsey & Company. https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america

MCTI. (2023a, setembro). *Ministério anuncia mais quatro Centros de Pesquisa em Inteligência Artificial*. Ministério da Ciência, Tecnologia e Inovação. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2023/09/ministerio-anuncia-mais-quatro-centros-de-pesquisa-em-inteligencia-artificial

MCTI. (2023b, dezembro). *MCTI anuncia revisão da Estratégia Brasileira de Inteligência Artificial*. Ministério da Ciência, Tecnologia e Inovação. https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2023/12/mcti-anuncia-revisao-da-estrategia-brasileira-de-inteligencia-artificial

MGI. (2025a). *Estratégia de Inteligência Artificial para o Setor Público Federal*. Ministério da Gestão e da Inovação em Serviços Públicos. https://www.gov.br/gestao/pt-br

MGI. (2025b, setembro). *Na Soberania Digital: Brasil consolida a Nuvem de Governo como passo para a independência e a proteção de dados*. https://www.gov.br/gestao/pt-br/assuntos/noticias/2025/se/na-soberania-digital-brasil-consolida-a-nuvem-de-governo-como-passo-para-a-independencia-e-a-protecao-de-dados

National University. (2025). *59 AI Job Statistics: Future of U.S. Jobs*. National University. https://www.nu.edu/blog/ai-job-statistics/

NIC.br. (2025). *Observatório Brasileiro de Inteligência Artificial (OBIA)*. Núcleo de Informação e Coordenação do Ponto BR. https://www.nic.br/

NIST. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

NUPEPA. (2025). *Inteligência Artificial nas Ciências Sociais e Humanas: Usos, reflexões e impacto* (Observatório Social para a Inteligência Artificial e Dados Digitais (OSIADD)). Núcleo de Pesquisa em Políticas Públicas e Ação Coletiva. https://zenodo.org/records/15257282

Observatório da Imprensa. (2024). Acesso à internet no Brasil reproduz desigualdades e exige políticas públicas. *Observatório da Imprensa*. https://www.observatoriodaimprensa.com.br/digital/acesso-a-internet-no-brasil-reproduz-desigualdades-e-exige-politicas-publicas/

OECD. (2024). *OECD Principles on Artificial Intelligence (atualização)*. Organização para a Cooperação e Desenvolvimento Econômico. https://oecd.ai

Olga, A., Tzirides, Saini, A., Zapata, G., Searsmith, D., Cope, B., Kalantzis, M., Castro, V., Kourkoulou, T., Jones, J., da Silva, R. A., Whiting, J., & Kastania, N. P. (2023). *Generative AI: Implications and Applications for Education* (https://scispace.com/pdf/generative-ai-implications-and-applications-for-education-2bzkc2ap.pdf; Versão 3). arXiv. https://doi.org/10.48550/ARXIV.2305.07605

Penso. (2025, agosto). *Soberania de Dados, CLOUD Act e Riscos Jurídicos: O alerta que o Brasil precisa ouvir*. Penso Consultoria. https://www.penso.com.br/soberania-de-dados-cloud-act-e-riscos-juridicos-o-alerta-que-o-brasil-precisa-ouvir/

Pires, R., Abonizio, H., Almeida, T. S., & Nogueira, R. (2023). Sabiá: Portuguese Large Language Models. *Brazilian Conference on Intelligent Systems (BRACIS)*, 226–240. https://doi.org/10.1007/978-3-031-45392-2_15

Pretto, N. D. L. (2024, maio). *20 Anos do Software Público Brasileiro: Pioneirismo Nacional em Bens Públicos Digitais*. https://movimento.softwarelivre.tec.br/noticias/20-anos-do-software-publico-brasileiro-pioneirismo-nacional-em-bens-publicos-digitais/

ProPublica. (2025). *About Us*. https://www.propublica.org/about/

PwC Brasil. (2025). *Casos públicos de uso de IA no setor de saúde*. PwC. https://www.pwc.com.br/pt/estudos/setores-atividade/saude/2025/casos-publicos-de-uso-de-ia-no-setor-de-saude.html

Queiroz, [Nome completo não especificado]. (2024). *Soberania digital e dependência de infraestruturas tecnológicas estrangeiras*.

Raulin, M. L. F., & Angel, D. J. (2025). INTELIGÊNCIA ARTIFICIAL NA MEDICINA: IMPACTOS E DESAFIOS. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, *11*(1), 2801–2814. https://doi.org/10.51891/rease.v11i1.18024

Rede pela Soberania Digital. (2024). *Carta pública ao presidente Lula: Plano Digital para a Soberania Nacional*.

Reuters. (2024, julho). *Brazil proposes $4 billion AI investment plan*. Reuters. https://www.reuters.com/technology/artificial-intelligence/brazil-proposes-4-billion-ai-investment-plan-2024-07-30/

Santos, D. L. dos. (2023). *Inteligência artificial aplicada à educação: Transformação ou desintegração da escola?* [PhD Thesis, Universidade de São Paulo]. https://doi.org/10.11606/T.48.2023.tde-21112023-155836

SBPC. (2024). *Povos tradicionais e biodiversidade: Conhecimentos ancestrais complementares à IA para sustentabilidade*. Sociedade Brasileira para o Progresso da Ciência. https://portal.sbpcnet.org.br/

SERPRO. (2024, novembro). *Inteligência Artificial desenvolvida pelo Serpro vai transformar o Carf*. Serviço Federal de Processamento de Dados. https://www.serpro.gov.br/menu/noticias/noticias-2024/carf-ia-serpro

Silva, A. P. da, Feferbaum, M., Camelo, A. P., Silveira, A. C. R. D., Katano, B. Y. S., Wagner, F. R., Gatto, R. F., Lana, P. de P., & Schippers, L.-M. (2024). *Relatório de pesquisa: Soberania digital: Para quê e para quem? Análise conceitual e política do conceito a partir do contexto brasileiro*. Centro de Ensino e Pesquisa em Inovação (CEPI), FGV Direito SP. https://hdl.handle.net/10438/35312

Stanford HAI. (2025). *AI Index Report 2025*. Stanford University. https://aiindex.stanford.edu

Syuntyurenko, O. V. (2022). Predicting Potential Threats and Megarisks in Information Technology Development. *Scientific and Technical Information Processing*, *49*(1), 48–59. https://doi.org/10.3103/S0147688222010130

Thakre, Prof. N., Niveditha, Prof. K. P., Jain, Prof. P., Soni, D., & Vishwakarma, M. (2023). Exploring the Ethical Dimensions of AI in Decision-Making Processes. *International Journal of Innovative Research in Computer and Communication Engineering*, *11*(10), 11325–11331. https://doi.org/10.15680/IJIRCCE.2023.1110048

UK government. (2025, fevereiro). *The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023*. GOV.UK.

https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023

UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence*. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org

UNESCO. (2022). *Recomendação sobre a Ética da Inteligência Artificial*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137_por

United States. (2022). *What is the Blueprint for an AI Bill of Rights?* The White House. https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/what-is-the-blueprint-for-an-ai-bill-of-rights/

United States. (2023, outubro). *Executive Order 14110—Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*. The White House. https://www.presidency.ucsb.edu/documents/executive-order-14110-safe-secure-and-trustworthy-development-and-use-artificial

United States. (2025, janeiro 23). *Removendo Barreiras à Liderança Americana em Inteligência Artificial*. The White House. https://www-whitehouse-gov.translate.goog/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt&_x_tr_pto=tc

Valente, G. (2025, junho). Brazil bets big on AI infrastructure. Em *Infralogic*. ION Analytics. https://ionanalytics.com/insights/infralogic/brazil-bets-big-on-ai-infrastructure/

Vujović, D. (2024). Generative AI: Riding the new general purpose technology storm. *Ekonomika Preduzeca*, *72*(1–2), 125–136. https://doi.org/10.5937/EKOPRE2402125V

Yale Budget Lab. (2025). *Evaluating the Impact of AI on the Labor Market: Current State of Affairs*. Yale Budget Lab. https://budgetlab.yale.edu/research/evaluating-impact-ai-labor-market-current-state-affairs