KEY THEMES IN AI REGULATION:

The local, regional, and global in the pursuit of regulatory interoperability

INSTITUTIONAL SUPPORT

DataPrivacyBR Research

HEINRICH BÖLL STIFTUNG
RIO DE JANEIRO

Technical Data Sheet

Data Privacy Brazil is a non-profit civil society organization that promotes the protection of personal data and other fundamental rights in the face of the emerging new technologies, social inequalities, and power asymmetries. It comprises a multidisciplinary team from different regions of Brazil that conducts research of public interest, produces technical notes, and analytical texts on emerging issues, and provides training for decision-makers and society at large.

The Association believes that the protection of personal data is one of the cornerstones of democracy and must be approached from the perspective of social justice and power asymmetries. Therefore, we are committed to encouraging a culture of data protection and ensuring that digital rights are fundamental rights for everyone, conducting research that is open to the public, guided by a strong social commitment, and ethically funded.

For more information about the organization, the impact of its projects, and how research is supported, please visit www.dataprivacybr.org.

Translation Beatriz Nunes

Directors

Bruno Bioni, Rafael Zanatta & Mariana Rielli

Coordinators

Carla Rodrigues, Jaqueline Pigatto, Pedro Saliba & Victor Barcellos

Advocacy

Vinícius Silva

Researchers

Eduardo Mendonça, Gabriela Vergili, Júlia Mendonça, Horrara Moreira, Louise Karczeski, Marina Meira, Paula Guedes & Nathan Paschoalini

Communication

Alicia Lobato, João Paulo Vicente, Rafael Guimarães, Rafael Regatieri & Roberto Junior

Management & Finance

Elisa Bayón & Matheus Arcanjo

Press

For clarifications about the document and interviews, please contact the Association via email at imprensa@dataprivacybr.org

How to cite this document

BIONI, Bruno; GARROTE, Marina; GUEDES, Paula. *Key Themes in AI Regulation: The local, regional, and global in the pursuit of regulatory interoperability*. São Paulo: Data Privacy Brazil Research Association, 2023.

Table of Contents

KEY THEMES IN AI REGULATION: THE LOCAL, REGIONAL, AND GLOBAL IN THE PURSUIT OF REGULATORY INTEROPERABILITY1[1]

Bruno Bioni[2]

Marina Garrote[3;4]

Paula Guedes[5]

1 This publication is the result of the project Where the sabIA sings: artificial intelligence governance and regulation in Brazil, funded by the Luminate Foundation, Eko, and the Heinrich Böll Stiftung. For more information, visit: https://www.dataprivacybr.org/projeto/onde-canta-o-sabia-governanca-e-regulacao-de-inteligencia-artificial-a-partir-do-brasil/.

2 PhD in Commercial Law and Master's Degree in Civil Law from the University of São Paulo Law School. He was a study visitor at the Personal Data Protection Department of the European Data Protection Board/EDPB and the Council of Europe, and a visiting researcher at the Law, Technology and Society Research Center of the University of Ottawa Faculty of Law. He is the author of the book Personal Data Protection: the role and limits of consent. He is a member of the Latin American Network of Studies on Surveillance, Technology and Society/LAVITS, and also of the International Association of Privacy Professionals – IAPP, with CIPP/E Certification. He is the founding director of Data Privacy Brazil, an intersection between a course school and a research association in the field of privacy and data protection..

3 Lawyer. Master's student in Law at the New York University School of Law. Master's degree in Civil Procedure from the University of São Paulo. Specialist in Gender and Sexuality from the Latin American Center on Sexuality and Human Rights at the Institute of Social Medicine of the State University of Rio de Janeiro.

4 Marina Garrote, co-author of this publication, engaged in the structuring and methodology of the research as a whole and contributed with writing and research during the months of May and June 2023, especially on axes 1.

5 Lawyer. PhD student in Law and Artificial Intelligence at the Portuguese Catholic University – Porto Regional Center and a Master's Degree in International and European Law from the same institution; specialist in Digital Law at ITS-Rio in partnership with UERJ. Member of the Legalite Law and Technology Research Center at PUC-Rio. Researcher at the Data Privacy Brasil Research Association.

1. Executive Summary and Methodological Notes

The primary objective of this position paper is to organize basic concepts and theoretical frameworks on three structural themes of any regulatory proposal on artificial intelligence (AI), examining how they have been covered by legislative initiatives in Brazil. Particularly, that of Bill 2338/2023, comparing it with laws, regulatory projects, and soft law documents[6] from other countries and international entities.

This is not an exhaustive work that will cover all the issues arising in each of the themes, but the central purpose is to inform interested parties about the current state of the art in terms of AI regulation, especially during the legislative process in Brazil.

In addition, the secondary objective of this study is to map the level of convergence of Brazilian proposals with those of other countries and multilateral and international organizations. A qualitative analysis that captures the regulatory rationale for intersecting with a global movement in AI governance, while at the same time not losing sight of the nuances of what is happening in Brazil (especially with regards to Bill 2338/2023). Ultimately, the reader will have guidelines to assess the extent to which the Brazilian discussion is interoperable[7] and to understand its particularities, diverging from uncritical legal transplantation and even colonization[8] compared to discussions outside the country.

The main conclusions of this study, which can assist in organizing the regulatory debate on the subject, are:

6 Soft law refers to rules of conduct with normative content, however, lacking formal binding force, and therefore generating practical effects on the behavior of individuals and institutions through self-regulation by private actors (DA SILVA, Paula Guedes Fernandes. Artificial Intelligence in the European Union: ways to regulate the technology that already regulates us. In: MENDES, Gilmar Ferreira; DE MORAIS, Carlos Blanco. Governance of the Legal Order in Transformation. Proceedings of the X Lisbon Legal Forum, 2022, p. 589). Available at: https://www.forumjuridicodelisboa.com/2023-anais; TRUBEK, David M.; COTRELL, Patrick; NANCE, Mark. "Soft Law," "Hard Law," and European Integration: Toward a Theory of Hybridity. Legal Studies Research Paper Series, Winsconsin, n. 1002, p. 1-42, Nov. 2005. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=855447.

7 Colin Bennett, in his studies on the field of personal data protection, discusses "regulatory convergence" in this area. In other words, although the law is traditionally recognized to vary from country to country, there are certain scenarios in which it is possible to observe regulatory standardization when there are national regulations with similar foundational elements. BENNETT, Colin J.; RAAB, Charles D., Revisiting the governance of privacy: Contemporary policy instruments in global perspective. Regulation & Governance, Vol. 14, Issue 3, p. 447-464, 2018.

8 The formation of legal culture in Latin America, especially in Brazil, has been characterized by the often uncritical importation of foreign regulations, primarily from the European Union, into our reality. This situation results in certain regulatory legal provisions not being based on Brazilian needs as a country in the Global South, but rather as a mere importation of systems that cater to the interests and needs of a society different from ours in many respects; FERRAZZO, Débora; DUARTE, Francisco Carlos. Colonial Influence in Latin American Law. Available at: www.publicadireito.com.br/artigos/?cod=f376b8ae6217d18c.

(1) **How to regulate and navigate between the general and the sectoral:** the continuous emergence of new regulations directed at AI, whether through bills/regulations or international documents from globally relevant actors, reveals the global trend where the discussion is no longer about whether **to** regulate the use of this technology, but **how** to regulate it. Due to the continued production of negative externalities across sectors, sectoral regulations are not sufficient. However, this does not negate the need for a governance arrangement that navigates between the general and the specific precisely to translate general governance norms into the particularities of a given context. In other words, a general law does not exclude, but rather, opens space for sectoral regulation to flourish based on common foundations across different sectors of the economy;

(2) **Responsible and Socioeconomically Resilient Innovation:** the aim should not be to seek any kind of technological progress, but one that is socioeconomically responsible. **The trade-off is not between innovation and the protection of fundamental rights and freedoms**, but rather about what type of innovation: **whether it reinforces or undermines the democratic rule of law**. Therefore, the term "innovation" has been adjectived as responsible. From this premise, regulatory proposals - especially those that are cross-cutting rather than sectoral - have the potential to catalyze technological, economic, and social development. This is particularly true with the emergence of so-called foundation AIs (e.g., generative) that will have various uses in diverse contexts (downstream applications);

(3) **Flexible regulatory target and a dynamically balanced regulation (asymmetric risk-based regulation):** due to its regulatory object spreading across various sectors and contexts, it's impossible to have a homogeneous response. For this reason, a significant common point observed among a wide range of options is the model of asymmetric risk-based regulation. The idea is to calibrate the weight of regulation - the intensity of obligations, rights, and responsibilities of a particular regulated entity - according to the level of risk in a given context. This means that regulatory efforts and governance obligations are not the same for all cases, even within the same sector, or even for all actors in the AI chain. This regulatory decision gained prominence with the proposal for AI regulation from the European Union, but it is already present in various other sources, coming from the OECD, UNESCO, Canada, the Council of the European Union, and even the Unit-

ed States. This approach is seen as positive for stimulating innovation, as it proportionally doses the degree of regulatory intervention according to the level of risk, thus not creating an excessive burden of obligations. In the Brazilian context, the bill that most aligns with the international trend is Bill 2338/2023, as unlike others, it seeks to minimally proceduralize a both dynamic and balanced classification according to the contextual risk of AI

(4) **There are several models of risk regulation:** the risk-based regulatory technique is not monolithic; on the contrary, it has several variations and even extremes that range from state monopoly (command and control regulation) to private (self-regulation) in the task of risk management. Furthermore, it is possible to have hybrid models, as is the case with the vast majority of AI regulation proposals, which rely on a co-regulation model where there is allocation of state resources and incentives for economic agents themselves to come together in a kind of public-private partnership. However, even in such hybrid models, there are important nuances, such as a model with greater "democratic oversight," where risk is subject to greater public scrutiny and social control. The best example of this, and with particular relevance to the Brazilian legal culture, is the environmental field. In this context, there are different ways in which civil Society collaborates, such as participating in the development and execution of environmental policies. This can occur through the involvement of civil Society representatives in collegiate bodies with normative powers, or through the opportunity to participate in public hearings as part of environmental impact studies, or even by participating in municipal environmental councils. Ultimately, the major tensions of a risk-based regulation model lie precisely in a method of greater or lesser social porosity, whose distortions have historically been observed in various regulated sectors. While AI regulation may exacerbate these asymmetries and risk regulation may be less democratic and more technocratic, it can, on the other hand, paradoxically serve as a window of opportunity for the purpose of equalization and, consequently, greater legitimacy in regulatory production with an increased social engagement. In Brazil, Bill 2338/23 advances toward a model of democratic risk oversight, but it could be improved by, for example, including a more programmatic chapter in this regard alongside existing provisions, which can be reinforced, for public participation in the assessment, classification, and management of risks associated with AI.

(5) **The different levels of risk:** the asymmetric risk regulation model is made

up of different categories, which can vary according to the chosen methodology. Internationally, there is a tendency to define what would be unacceptable and high risks, leaving the others (low and medium) as residual, as is the case with the EU AI Act and Chile's Bill 15869-19. The nomenclature for each of these risk levels may vary.

Metaphorically speaking, the risk gradation can be associated with the image of a pyramid, where the base represents cases of lower risk (with fewer obligations), the middle represents high risks (with obligations imposed to permit the implementation of technology), and the top represents excessive/extremely high/unacceptable risks. In this last case, it involves significant regulatory intervention by preventing the use of technology because it is understood to bring more risks than benefits. To define each of these levels, it is important to establish qualitative elements. In other words, instead of just defining the degrees of risk (for example, low/medium/high risk) in general terms, it is indispensable to have minimum criteria for identifying systems at each of these levels.

In Brazil's case, Bill 2338/23 is the only one to establish such a division by creating the categories of excessive and high risk, as well as a residual category for systems not classified by the first two levels. Each of these categorizations will trigger different obligations, which may be more or less intense, thus also gauging regulatory resources.

(6) **Risk as a dynamic element:** in addition to establishing risk levels, it is essential to also define minimum criteria for identifying systems at each of these levels, based on qualitative and quantitative elements. From international experience, examples of these criteria could include: context, scope, level of automation, degree of explainability, potential impact on individuals, quantity of data processed, among others. The definition of criteria provides legal certainty for regulated entities by avoiding excessive regulatory generalization. There is a certain pattern of directing the highest regulatory burden towards high-risk AI systems, which, although not prohibited ex-ante, must comply with a range of obligations for their development and use. They are generally presented by a list of illustrative cases that are complemented by qualitative and quantitative criteria for updating these hypotheses, as is the case in Bill 2338/23 (Brazil),

AIDA (Canada), AI Act (European Union), and Bill 15869-19 (Chile). This provision of criteria for updating cases of high-risk AI allows for legislation to be kept alive and not subject to the passage of time.

(7) **The difficult conciliation of regulation based on risk and on rights - risk taxonomy as one of the possible indicators (proxy):** as with AIDA (Canada), the AI Act (European Union), Bill 15869-19 (Chile), and the OECD and UNESCO guidelines, Bill 2338/23 (Brazil) is adamant that a risk-based regulation should be used to reinforce and not undermine the rights of people and groups affected by AI. To this end, the Brazilian bill opted to systematize these rights and, correspondingly, the related duties. As a result, there is a topographical structure that is consistent with this promise of harmonization. Especially with the provision that some rights apply regardless of the AIs system risk. Another indicator of this possible conciliation is how such regulatory proposals incorporate the taxonomy of excessive (unacceptable) and high risks. While some proposals have opted, for example, for the exclusion of biometric data for law enforcement purposes, others lean towards a moratorium until such practice is regulated. Furthermore, there is a significant variation in the illustrative list of high-risk AIs, as well as in the quantitative and qualitative criteria for a dynamic taxonomy in this regard. In other words, the regulatory appetite for intervention is not the same when foreseeing ex ante in which contexts certain rights and freedoms are non-negotiable, as well as the expansion of situations in which the regulatory burden would be excessive for the protection of individuals or affected groups. This intertwining of risk classification logic with rights serves as a potential indicator of the aforementioned conciliation.

(8) **Public, inclusive, and rights-based Algorithmic Impact Assessments (AIA), focusing not only on individual rights but also on social rights:** to be an effective accountability tool, various regulatory proposals identified have opted for a minimal proceduralization of AIA based on a threefold approach. The initial aspect is that of transparency – at least one version that can be disclosed – to ensure that the outcomes of risk management assessments are shared with the entire Society, thus becoming an oversight agent as well. In fact, some of the regulatory proposals, such as those from Brazil, Europe, and the United States, envision the establishment of a public database on high-risk AI systems. The second aspect is significant multi-sectoral public participation from individuals and communities potentially affected,

especially from the most vulnerable and marginalized, ensuring that the assessment process and final outcome are as fair and accurate as possible for the reality in which it will be applied. This fosters greater legitimacy and democratic oversight of AI systems, moving towards multi-participatory governance and co-generation of technology throughout its lifecycle. Given Brazil's established history of multi-sectoral governance on the Internet and the affirmative provisions of Bill 2338/23 regarding the right and duty of co-deliberation on acceptable AI risks, Brazil could potentially have successful regulatory experiences in the future. The third aspect is the variety of risks and benefits to be evaluated, as there is still a predominance of adverse effects on individual fundamental rights compared to social rights. Unlike UNESCO's recommendations, Bill 2338/23 does not expressly include diffuse and collective rights, such as the right to work and to the environment as part of its framework.

(9) **Regulation that is in line with local socio-technical and economic aspects:** as Brazil and the countries of the global south must draw up national strategies that adhere to local challenges and opportunities that are different from those of the global north, it is imperative that AI regulation is not an acritical copy of regulatory models from other contexts. Among the bills currently pending in Congress, only Bill 2338/2023 makes progress in this type of tropicalization of the debate. The bill acknowledges that Brazil is permeated by asymmetries and structural inequalities by including, among the definitions in Article 4, the concepts of direct and indirect discrimination from the Inter-American Convention against Racism, which Brazil endorsed in 2022 with constitutional status. In addition, it provides a normative rationale for the participation of vulnerable groups in the assessment and management of AI risks that affect them, as well as specific rules for the implementation of AI in the public sector, given the scenario that the most socio-economically vulnerable will be the most positively or negatively affected. However, the bill is still rather reactive in not providing, as other regulatory frameworks have done (e.g., Brazilian Internet Bill of Rights and General Data Protection Law), for a more programmatic chapter combined with multisectoral institutional arrangements.

(10) **Generative AI and stress testing of AI regulation proposals:** in November 2022, with the release of Chat GPT by OpenAI, the discussion on generative AIs came to the fore. Despite an alarmist narrative that their disruption

would prevent regulation by existing laws and regulatory proposals, the truth is that these AIs also need to and can be governed, with the appropriate adaptations. One of the first challenges faced is its very definition, since there are different nomenclatures used, such as generative AI, foundation models, large language models, large generative AI models, among others. However, due to common characteristics, it is possible to equate them since the considerations regarding their regulation are similar.

The second and perhaps greatest challenge of generative AIs is that, because they serve different purposes (which are not always predictable), they put pressure on the risk-based regulatory model, which is currently predominant in the AI field, given that it is inherently contextual. In an attempt to mitigate this challenge, it is possible to include, within the risk model, both the idea of "general purpose AI" and the risk analysis that can reasonably be expected, along with those that are known and predictable. As a result, even if we are not fully aware of the existence of some risks, preventive measures can be taken, as is already the case in the current text of Bill 2338/2023 and in the European Parliament's version of the EU AI Act. Another solution is to better develop the actors involved in the production chain of Generative AI systems so that the obligations of each of them can be broken down - as the latest version of the European proposal has done by calling for cooperation between these agents.

2. The Artificial Intelligence Regulatory Context in Brazil

In line with international processes, Brazil has been moving towards AI governance for some years now. In April 2021, the Brazilian Artificial Intelligence Strategy (EBIA) was published by the Ministry of Science, Technology, and Innovations (MCTI), which established nine central thematic axes for the development of AI systems in Brazil. Among them, the "Legislation, regulation, and ethical use" axes stands out, aiming to seek a balance between the protection of fundamental rights, technological development, and the creation of legal parameters to establish legal certainty regarding the responsibility of the agents involved in the AI value chain[9].

However, EBIA has received a lot of criticism because of its generic approach and lack of planning by, for example: a) not indicating the actors responsible for governance; b) not deepening the analysis of applicable methods for critical issues (such as transparency and explainability); and (iii) not critically reflecting on the use of AI in highly risky contexts, such as public safety[10].

Concurrently with EBIA, since before 2021, there have been several bills filed in the Brazilian Congress regarding AI regulation. Among them was Bill 21/2020, presented expressly as an attempt to implement legislation on the use of AI systems in Brazil[11] and whose urgency and approval were deliberated in the Chamber of Deputies in the same year.

The final text of Bill 21/20 approved by the Chamber of Deputies in September 2021, without exhausting all mechanisms of significant public participation[12], established general principles for the development and application of AI in Brazil. It brought only a few guidelines specifically for the public sector, maintaining the sectoral self-regulation model of the technology, while also not providing a list of rights and responsibilities. Furthermore, the approved text lacked normative density, the implementation of effective governance tools, and the addressing of specific risks that the development and use

9 Ministry of Science, Technology, and Innovations (MCTI). Brazilian Artificial Intelligence Strategy – EBIA. July 2021. Available at: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-documento_referencia_4-979_2021.pdf. p. 16.

10 GASPAR; Walter B.; DE MENDONÇA, Yasmin Curzi. Artificial Intelligence in Brazil still needs a strategy. A report by the Center for Technology and Society at FGV Direito Rio. May 2021. Available at: https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/30500/EBIA%20pt-br.pdf?sequence=3&isAllowed=y.

11 Data Privacy Brazil Research Association. Technical Note – Data Privacy Brazil's contributions to Bill 21, of February 4, 2020. 2021. Available at: https://www.dataprivacybr.org/wp-content/uploads/2021/09/dpbr_notatecnica_pl21.pdf. p. 3.

12 In episode 78 of the Democracy Podcast, Bruno Bioni points out that the process of approving Bill 21/20 in the Chamber of Deputies in 2021 could have made use of other public participation tools available, other than just public hearings, such as public consultations, which would have given space for collaboration by people and groups who did not have a voice during the hearings held, giving greater legitimacy to the legislative process. Data Privacy Brazil. Dadocracia – Ep. 78 – Legal Framework for AI. Dadocracia, published in Nov. 2021. Available at: https://open.spotify.com/episode/15BWzRa4cWVRoOjtGGPm4T?si=v7X-iVnWQ3eeIArlGmKaUg.

of AI in Brazil could trigger[13].

Following the approval of the urgent procedure, Bill 21/20 was subject to criticism, especially from the academic community and civil society, which led to a social mobilization calling for greater debate and public participation on the proposal[14]. The text, if approved, would end up leaving Brazil out of step with the international regulatory framework[15].

In this context, in February 2022, Senator Rodrigo Pacheco, president of the Federal Senate, set up a Commission of Jurists responsible for subsidizing the drafting of a substitutive on artificial intelligence in Brazil (CJSUBIA). The Commission spent 240 days working hard on the drafting of the bill. In addition to international seminars and public hearings with more than 90 (ninety) people heard, a public consultation was also opened to allow any individual or entity to contribute to the debate[16].

In December of the same year, the Final Report of the activities of the CJSUBIA was published, comprising over 900 pages, which included, in addition to the history of its activities and the public participation processes externalized in the written contributions, public hearings and international seminar, the draft substitute to Bills 5.051/2019, 21/2020 and 872/2021.

Consisting of 45 articles, the new draft aims to demystify the supposed trade-off between a regulation that guarantees rights and economic development and innovation, by establishing a risk-based and rights-based approach through asymmetric regulation,

13 DA SILVA, Paula Guedes Fernandes; GARROTE, Marina Gonçalves. Insufficiency of ethical principles to standardize Artificial Intelligence: anti-racism and anti-discrimination as vectors of AI regulation in Brazil. POLITICS, September 2022. Available at: https://politics.org.br/edicoes/insufici%C3%AAncia-dos-princ%C3%ADpios-%C3%A9ticos-para-normatiza%C3%A7%C3%A3o-da-intelig%C3%AAncia-artificial-o.; Data Privacy Brasil. Dadocracia – Ep. 78 – AI Legal Framework. Dadocracia, published in Nov. 2021. Available at: https://open.spotify.com/episode/15BWzRa4cW-VRoOjtGGPm4T?si=v7X-iVnWQ3eeIArlGmKaUg; Data Privacy Brazil. Dadocracia – Ep. 80 – More AI Legal Framework. Dadocracia, published in Dec. 2021. Available at: https://open.spotify.com/episode/0t4RrO7EwIjrdpmvzht79Z?si=OiOyUXcOT5-kHOnr6qhzKA&nd=1; *Estadão*. "Brazil's most important technology law is not under debate," says expert. Bruno Romani, published on Dec. 7th, 2021. Available at: https://www.estadao.com.br/link/cultura-digital/mais-importante-lei-de-tecnologia-no-brasil-nao-esta-sendo-debatida-diz-especialista/. *Folha de São Paulo*. Brazil rushes through artificial intelligence law, say experts. Amanda Lemos, published on July 18, 2021. Available at: https://www1.folha.uol.com.br/mercado/2021/07/brasil-apressa-lei-para-inteligencia-artificial-dizem-especialistas.shtml.

14 Rights on the Net Coalition. Artificial Intelligence cannot be regulated at the drop of a hat. Published on September 23, 2021. Available at: https://direitosnarede.org.br/2021/09/23/inteligencia-artificial-nao-pode-ser-regulada-a-toque-de-caixa/; Rights on the Net Coalition. Brazil is not ready to regulate artificial intelligence. Published on December 7th, 2023. Available at: https://direitosnarede.org.br/2021/12/07/brasil-nao-esta-pronto-para-regular-inteligencia-artificial/.

15 Data Privacy Brazil Research. Technical Note – Data Privacy Brazil Contributions to Bill 21, of February 4, 2020. Available at: https://www.dataprivacybr.org/wp-content/uploads/2021/09/dpbr_notatecnica_pl21.pdf; Da Silva, Paula Guedes Fernandes; Garrote, Marina. Insufficiency of ethical principles to standardize Artificial Intelligence: anti-racism and anti-discrimination as vectors of AI regulation in Brazil. PoliTICS, set. 2022. Available at: https://politics.org.br/edicoes/insufici%C3%AAncia-dos-princ%C3%ADpios-%C3%A9ticos-para-normatiza%C3%A7%C3%A3o-da-intelig%C3%AAncia-artificial-o.

16 For more information on the activities carried out by CJSUBIA, see: https://legis.senado.leg.br/comissoes/comissao?codcol=2504.

i.e., increasing the burden on regulated agents according to the level of risk of their AI system. According to the explanatory memorandum:

> "Its normative objective is to combine a risk-based approach with a rights-based regulatory model. While governance instruments are provided to ensure accountability and reward the good faith of economic agents who effectively manage the risks surrounding the design and implementation of artificial intelligence systems, there is also a strong burden of obligation for the flourishing of individual and social scrutiny in relation to its use.[17]"

In May 2023, the preliminary draft bill (APL) was converted by the president of the chamber into a new bill, numbered 2338/2023. Currently, said bill is under analysis in the Temporary Internal Committee on Artificial Intelligence in Brazil (CTIA), recently established within the scope of the Federal Senate to examine, within 120 days, the aforementioned bill, as well as any new projects addressing AI related matters.

Therefore, although there are other bills subject to analysis, this report, when referring to the Brazilian context of AI regulation, will focus on the projects that have been more widely discussed publicly, with a significant emphasis on Bill 2338/23 and Bill 21/20.

17 Commission of Jurists responsible for subsidizing the drafting of a substitutive on artificial intelligence in Brazil (CJSUBIA). Final Report. Federal Senate, Dec. 2022. p. 10 and 11.

3. Project Scope and Methodological Assumptions

The three main themes/axes chosen for analysis were:

(i) Risk-based regulation;

(ii) Algorithmic impact assessments (AIA); and

(iii) Generative AI

These themes were not chosen at random. Axes I and II were selected because they are central issues for a proper balance between risk-based and rights-based regulation. Item III represents one of the topics that has raised the most questions concerning the regulatory approach in recent times. Finally, a specific item has also been added to address the particularities of AI regulation in Brazil.

As stated previously, the primary objective is to map out the main discussions along the axes chosen in order to inform readers about the discussions currently taking place and thus provide a current state-of-the-art diagnosis in terms of regulating AI. To this end, a comparison will be made between the main legislative initiatives in Brazil and around the world so that the reader can visualize the possible choices to be made by Brazilian and international legislators, providing a critical perception of the range of options available.

When comparing the bills and other regulatory initiatives, charts will be used to highlight the similarities and differences between them. The goal is to provide a clearer picture of the implications that the different choices at stake are capable of producing, ultimately reducing the information asymmetry for those wishing to engage in this legislative and regulatory debate on AI. Additionally, it should be noted that most of the laws compared are from the global north or from international organizations (with the exception of the Chilean and the Brazilian proposals). The choice was intentional because these proposals are getting the most coverage in the Brazilian media. The goal is to determine whether there are significant nuances between them and between the Brazilian proposals, which could trigger a movement of normative colonization.

This research is only the first step. Subsequently, we aim to collaborate collectively with our peers from the global majority in comparative research in which they are protagonists-subjects and not just supporting actors-objects of analysis. We don't want to reproduce a common pattern in which the global South is only interviewed and analyzed but is not the author of intellectual productions[18].

18 In this regard, see the joint contribution of the Southern Alliance for the Global Digital Compact, whose authors are all entities from South America, Africa, and India; Southern Alliance for the Global Digital Compact: contribution for

For methodological purposes and for the reader's better understanding, it is necessary to clarify that, as the term "artificial intelligence" is broad and encompasses different technologies as an umbrella term, not all projects, laws, and documents compared will have exactly the same scope. For example, while some initiatives address AI, others focus solely on automated decisions. Similarly, some documents mention impact assessments focusing on different aspects (fundamental rights, democracy, rule of law, human rights, among others), while others refer to related instruments such as risk assessments.

the promotion of digital human rights.2023. Available at: https://www.dataprivacybr.org/documentos/southern-alliance-for-the-global-digital-compact/.

DOCUMENTS AND REGULATIONS THAT WILL UNDERGO ANALYSIS

Nacionais			
Regulation	Abbrevia–tion	Summary	Why compare and analyze?
Draft Bill 21/2020	Bill 21/20	The proposal was filed in the Chamber of Deputies by Congressman Eduardo Bismarck in 2020 to establish principles, rights, and duties for the use of artificial intelligence in Brazil. The bill was urgently approved in the Chamber of Deputies in September 2021 in the form of a substitutive by Congresswoman Luisa Canziani.	It was the first bill on AI in Brazil to make progress in Congress, bringing with it an initial set of principles and objectives and a soft-law approach, without any major governance instruments.
Draft Bill 2338/2023	PL 2338/23	The proposal was filed in the Federal Senate in May 2023 by Senator Rodrigo Pacheco. The bill's text is the result of months of work by a Committee of Jurists established aiming to provide input for the creation of a substitutive for bills on AI in Brazil pending analysis by the Senate, such as Bill 21/20.	Said bill was the result of 8 months (240 days) of work by a Committee of Jurists (CJSUBIA). During this period, the Committee enabled various forms of participation: public hearings, consultations, and an international seminar, all with contributions from experts in AI-related topics. All processes, stages, and studies conducted by CJSUBIA were synthesized in its final report delivered to the President of the Federal Senate in December 2022.
Draft Bill 759/2023	PL 759/23	The proposal was filed in the Chamber of Deputies in February 2023 by Congressman Lebrão with the intention of regulating AI systems in Brazil and creating an obligation for the Executive Branch to define a National Artificial Intelligence Policy.	This is yet another bill on AI in Brazil that is an example of a generalist and soft-law approach, which could be negative for the governance of AI systems in Brazil.
Draft Bill 872/21	PL 872/21	Proposal filed in 2021 by Senator Veneziano Vital do Rêgo to address the ethical frameworks and guidelines for the development and use of Artificial Intelligence in Brazil.	As with the previous bill, this is another example of a general bill aimed at regulating AI in Brazil.
Draft Bill 5051/19	PL 5051/19	Proposal filed in 2019 by Senator Styvenson Valentim to regulate the use of AI in Brazil.	Similarly, another bill whose central focus is to regulate the use of AI in Brazil and which takes a principled approach without effective governance measures.

General Data Protection Law (Law 13.709/2018)	LGPD	Brazilian law that regulates the processing of personal data by natural persons or legal entities, whether public or private.	It is a cross-cutting law, with a risk-based and rights-based approach, which bears a close resemblance to some of the regulatory techniques for the AI scenario.
Preliminary Draft Law on Data Protection for Public Security and Criminal Prosecution	CRIMINAL LGPD	The Preliminary Draft Law on Data Protection for Public Security and Criminal Investigation, prepared by a Committee of Jurists, established by the President of the Chamber of Deputies in November 2019. This preliminary draft sought to create a legal text to address the exception of application of the LGPD, which exempts its application for the processing of data for public security and criminal investigation. However, it is yet to become a bill, in the Brazilian Congress.	It is a draft law that provides regulatory tools similar to those proposed for AI, such as regulatory impact analysis and personal data protection impact reports. The preliminary draft law is an example of risk-based regulation specific to the State in the context of public security and criminal investigation. As it concerns the public sector, there is greater care in the proceduralization of impact assessments (regulatory and of the technological applications themselves) and in the processes of data disclosure.

Internationally Proposed Regulations				
Regulation / Organization	Abbreviation	Location	Summary	Why compare and analyze?
Proposal for a Regulation of the European Parliament on Artificial Intelligence (Artificial Intelligence Act)	EU AI Act	European Union	Regulation proposal created by the European Commission to regulate AI within the European Union. The first version was published in April 2021 and is currently in the final stages of discussion, pending approval by the Member States of the European Parliament (MEPs).	European Union Regulation Proposal for AI systems. It is a global reference in terms of risk-based AI regulation and will likely result in a new Brussels Effect[19].

19 The Brussels Effect refers to the unilateral ability of the European Union to regulate global markets by creating rules that raise standards worldwide. While these rules are not coercive for other regions of the globe, they end up becoming a global reference due to market forces, as multinational companies voluntarily extend these rules to govern their global operations. This has occurred in the fields of data protection, consumer health and safety, environmental protection, antitrust, and online hate speech; BRADFORD, Anu. The Brussels Effect: How the European Union Rules the World. Nova York: Columbia Law School, mar. 2020.

General Data Protection Regulation of the European Union	GDPR	European Union	A 2016 Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data within the European Union.	Although the scope of this regulation is the protection of personal data, it is an example of a regulation that can serve as inspiration for regulatory models for the AI scenario as well, since it presents a risk-based model and also provides for governance measures, such as the impact assessment of the protection of personal data.
Proyecto de Ley 15869/19	Chilean bill	Chile	Chilean bill introduced in the Chamber of Deputies on April 24, 2023, aiming to regulate artificial intelligence systems, robotics, and connected technologies in their different areas of application.	The bill is inspired by the EU AI Act, also determining a risk-based approach and a mandatory risk management plan for high-risk systems, although it has substantial differences, such as the need for prior authorization from the competent authority before AI systems can be applied. For this reason, it should be analyzed as another example of an AI regulation proposal inspired by the European model in the Latin American context.
NYC Bias Audit Law (Local Law 144)	NY Bias Audit	United States (Nova York)	New York State law, enacted in April 2023, which requires that a bias audit be performed on automated tools used for decision-making for employment purposes prior to their actual use.	A state law which represents an important initiative from the United States to regulate AI in the form of automated tools, with a focus on audits.
The Washington DC Algorithms Law (B25-0114)	Bill 25-114	United States (Washington)	The "Stop Algorithm Discrimination" bill was reintroduced in the District of Columbia in February 2023. The aim is to prohibit users of automated decisions from employing said decisions by means of a discriminatory eligibility criteria. Among the proposed obligations are mandatory annual audits and transparency requirements.	State bill, with specific concern for the discriminatory potential of automated decisions, which should be considered and thus, compared because of its particularity in dealing with cases of discrimination in this context.

Assembly Bill 331 on Automated Decision Tools	AB-331	United States (Califórnia)	The bill requires, among other things, that automated decision agents conduct an annual impact assessment for any decision tool.	A bill from the State of California aimed at the use of AI in automated decisions, which reinforces the importance that these tools undergo an annual impact assessment.
AI Disclosure Act of 2023 (federal USA)	-	United States (federal)	According to the proposal, all material generated by artificial intelligence technology would have to include an explicit notice that AI generated it.	In light of the discussions surrounding generative AIs, this US bill proposes greater transparency in the use of AI, which could serve as an example for other regulatory initiatives concerning AI.
Algorithmic Accountability Act EUA	AAA	United States (federal)	Bill reintroduced in the US Congress in February 2022. If approved, the bill will be binding and will oblige companies to evaluate the impact of automated systems in terms of bias and effectiveness.	The only federal bill in the US context that addresses the topic of AI, with a specific focus on ensuring that AI systems go through accountability mechanisms, such as conducting an impact assessment.
Canada's Artificial Intelligence and Data Act	AIDA	Canada	AIDA is part of Bill C-27, Digital Charter Implementation Act, 2022. AIDA represents a significant milestone in implementing the Digital Charter and ensuring that Canadians can trust digital technologies, guaranteeing that the design, development, and use of AI systems are safe and respect Canadian values.	The framework proposed by AIDA aims to be the first step towards a new regulatory regime created to guide innovation in AI in a positive direction, through coordination with other international initiatives. Therefore, the document serves as a summary of what has been implemented worldwide, as the Canadian goal was to bring this dialogue among foreign sources. The proposal moves towards a risk-based approach and introduces the impact assessment tool, while expressly providing for non-rivalry between regulation and innovation incentives.

Reports from *Ad Hoc* Committee on Artificial Intelligence from the Council of Europe	CAHAI	Council of Europe - *Ad Hoc* Committee on Artificial Intelligence (CAHAI)	The Committee was established within the framework of the Council of Europe, with a mandate from 2019 to 2021, to examine the feasibility and potential elements, based on extensive multilateral consultations, of a legal framework for the development, design, and application of artificial intelligence, based on the standards of the Council of Europe in the field of human rights, democracy, and the rule of law.	The CAHAI has been a source of various regulatory feasibility studies, including studies related to algorithmic impact assessment that considers Human Rights, Democracy, and the Rule of Law as guiding principles of analysis, which is essential for analyzing the potential structure of an algorithmic impact assessment.
Draft Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law	Convention	Council of Europe - Committee on Artificial Intelligence (CAI)	CAI has a mandate for 2022 and 2024 within the framework of the Council of Europe and its main deliverable is, by 15/11/2023, an appropriate legal instrument (Convention) for the development, design and application of artificial intelligence systems based on the Council of Europe's standards on human rights, democracy, and the rule of law, and conducive to innovation, in accordance with the relevant decisions of the Committee of Justices.	The forthcoming Council of Europe Convention on AI, Human Rights, Democracy, and the Rule of Law will be the first document to create binding rules to regulate AI at international level, the result of years of study and research by a working group specializing in the area. One of the main findings of the analysis is the choice of a risk-based regulation and the imposition of obligations to prepare impact assessments in certain cases.
Washington SB 5116 - 2021-22	SB 5116	United States (Washington)	It establishes guidelines for government use and purchase of automated decision-making systems, in the interests of protecting consumers, improving transparency, and creating more predictability in the market.	A state bill that provides for accountability tools to be implemented by public agencies seeking to develop, use or purchase AI systems for automated decision-making, including, for example, a mandatory "algorithmic accountability report," which should be submitted to the competent authority, which will then publish the document for public comment. This is an example of a specific project for public authorities that calls for the development of accountability tools with public participation.

Blueprint for an AI Bill of Rights	Blueprint	United States (federal)	A non-binding document published by the White House in October 2022 to guide the design, development, and deployment of AI systems. The document is based on 5 principles: (i) safe and effective systems; (ii) algorithmic discrimination protections; (iii) data privacy; (iv) notice and explanation; (v) human alternatives, consideration, and fallback.	Even though it doesn't have binding force, initially[20], it is an important reference document as it was created by the White House with the aim of guiding standards for the design, development, and implementation of AI systems in the United States.
Artificial Intelligence Risk Management Framework (AI RMF 1.0) - NIST	AI RMF (NIST)	United States (federal)	The AI RMF is a voluntary framework that seeks to provide organizations with a process to help address risks throughout the AI lifecycle, seeking to enable trustworthy and responsible/reliable AI systems. It is intended to help manage business and societal risks related to the design, development, deployment, evaluation, and use of AI systems. It was developed by the US Department of Commerce's National Institute of Standards and Technology (NIST). It has no binding force and is based on 4 main principles.	The goal is to provide resources for organizations that design, develop, deploy, or use AI systems to help manage their risks and promote the reliable and responsible development and use of AI systems. The framework is voluntary, not sector-specific, and irrespective of the size of the organization that intends to use it. It is an important practical model for AI risk management and is already being followed by different organizations, even though it is not binding at first[21]. It provides a well-defined methodology, including qualitative and quantitative criteria.

20 The document becomes binding for some institutions in certain cases, such as for federal government agencies after the publication of the Executive Order of October 30, 2023, by President Joe Biden (Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence) – see Seg. 10, 10.2 (b) (iv).

21 See footnote 14.

Directive on Automated Decision-Making + Algorithmic Impact Assessment tool	-	Canada	A directive applicable to automated decision-making systems developed or implemented after April 2020. This directive provides for a Risk Assessment Tool designed to help Canadian departments and agencies to better understand and manage the risks associated with automated decision systems. The tool is a questionnaire that determines the level of impact of an automated decision system (composed of 48 risk questions and 33 mitigation questions). It is a mandatory risk assessment tool designed to support the Treasury Board of Canada's directive on automated decision making.	The Canadian directive is an example of legislation already in force that specifically targets the use of AI in automated systems for making administrative decisions on granting social benefits. What sets it apart is the provision of a specific tool for public bodies to manage the risks of their systems, including the definition of impact levels and respective mitigation measures.
Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems	-	Canada	The code temporarily provides Canadian companies with guidelines for the responsible development and use of generative AI systems until a formal regulation comes into effect.	Although voluntary, it provides concrete best practices for the development and use of generative AI, which can serve as an example when considering the regulation of this use of an AI.
The Organization for Economic Co-operation and Development	OECD	International	The OECD supports governments by measuring and analyzing the economic and social impacts of AI applications to identify best practices for public policies, with a series of publications on AI and its governance. In 2019, the organization published Principles for AI and created an Observatory for AI public policies, in addition to having different studies on the subject, such as the model for classifying AI and a report on accountability in AI through governance and management of its risks, as well as the recent guide to interoperability between AI risk management systems.	The OECD principles represent the first AI policy model, serving as the basis for other national and international documents and for assessing the status of AI governance in each country. Therefore, the organization's documents, such as those aimed at assisting in the classification of AI systems and their accountability through the management of these risks, also serve as an important basis for regulatory models that intend to regulate the uses of this technology.

Interim Measures for the Management of Generative Artificial Intelligence Services	-	China	Rules adopted by China's Science and Technology Department for generative AI, effective as of August 15th, 2023.	It is the first Chinese document to deal with Generative AI tools, serving as an example of how the regulation of this technology can be developed.
Recommendation on the Ethics of Artificial Intelligence (UNESCO)	-	UNESCO (international)	Adopted by UNESCO in November 2021, it was the first global standard instrument for AI Ethics adopted by the 193 member states, with the protection of human rights and dignity as its guiding principle. In addition, the recommendations also include specific areas for political action, which help policymakers to translate fundamental values and principles into action. Finally, the recommendation already presents two practical methodologies that also help in its practical application: (i) Readiness Assessment Methodology (RAP); (ii) Ethical Impact Assessment (EIA).	The UNESCO Recommendation, as well as its practical tools (especially the EIA), are an international reference for countries and organizations wishing to develop, implement and use AI systems.
Executive Order on Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence	Executive Order	United States (federal)	The executive order, signed on October 30th, 2023, by President Joe Biden, has binding force for US public bodies, which will have to comply with a series of obligations regarding the protection of US citizens against the potential risks posed by AI systems.	The Executive Order makes the NIST risk assessment framework and Blueprint binding for the U.S. federal government, which makes their analysis necessary, especially Sec. 10, 10.2 (b) (iv), as well as suggestions for regulating foundation AI models (generative AI).

4. Thematic Axes of Analysis

AXES 1 – Risk-Based Regulation

In order to delve deeper into risk-based regulation, it is necessary to go back a few steps to explore the notion of risk, an element inherent to life that moves away from a dualistic meaning (of existing or not). Gellert (2017) conceptualizes risk as a tool that aids the decision-making process, directing its analysis not to its presumed existence, but to how much risk a given agent is able or willing to take on and how much they are able to mitigate[22].

Similarly, Hood et al (2001) define risk as a probability of adverse consequences, with risk regulation being a governmental interference in market or social processes to control these potential adverse consequences[23]. Citing Beck (1992), Hood et al note that human activity and technology in modernity that have the collateral effect of risks that depend on experts to assess and recognize, are collective, global and irreversible in their impact, resulting in a "risk society", unlike previous historical periods[24].

Risk can therefore be understood as "the ability to define what might happen in the future and choose between alternatives"[25], functioning as a tool for decision-making, insofar as it makes the uncertain certain. Its constituent elements are two distinct but linked operations: predicting the future (with the help of numbers) and making decisions based on this. Consequently, risk, although associated with something more quantifiable, can also be understood as a qualitative and evaluative element that needs to be assessed from different perspectives.

a.1) Asymmetric regulation and risk: an overview

Risk regulation, as Hood et al have shown, varies considerably in relation to which risks are chosen for regulation and the way in which regulation works, not only between different legal systems (different countries), but also within the **same** legal system[26].

---

22 GELLERT, Raphaël. Understanding the notion of risk in the General Data Protection Regulation. **Computer Law & Security Review**: The International Journal of Technology Law and Practice (2017). p. 02; GOMES, Maria Cecília O. Somewhere between method and complexity: understanding the concept of risk in the LGPD. *In*: **Current data protection issues**. PALHARES, Felipe (Coord.). São Paulo: Thomson Reuters Brasil, 2020, pp 245-271..

23 HOOD, Christopher; ROTHSTEIN, Henry; BALDWIN, Robert.The Governance of Risk: Understanding Risk Regulation Regimes. Nova York: **Oxford University Press**, 2001. ISBN 0-19-924363-8.

24 In the authors' words: *"(...) we live today in a 'risk society'. By that Beck means that risk has a different significance for everyday life from that applying in previous historical eras. Human activity and technology in 'advanced modernity',he claims, produces a side-effect risks that need specialized expertise to assess and recognize, are collective, global, and irreversible in their impact, and thus potentially catastrophic on a scale never seen before."*; HOOD, *et al*, 2001, p. 3.

25 BERNSTEIN, Peter L. Against the Gods: The Remarkable Story of Risk. Wiley, 1996. p. 2.

26 HOOD *et al*, 2001.

With regard to the Artificial Intelligence systems' governance, Kaminski (2022) notes the choice of risk regulation tools in both the United States and Europe. This common choice, however, does not mean that there is a single risk regulation model in the regulations analyzed. The author highlights four risk regulation models. They are:

> (i) quantitative risk regulation model, which originated in US administrative law;
>
> (ii) risk regulation model that establishes democratic oversight for problems that affect the entire population, such as US environmental legislation (NEPA);
>
> (iii) risk regulation model that distributes the application of legislation and regulatory capacity based on risks, i.e., allocates regulatory resources contextually and transversally, as is the case in the United Kingdom;
>
> (iv) corporate risk regulation model.
>
> (KAMINSKI 2022)

The risk regulation model, which originated in US administrative law in the 1960s and 1980s to regulate health, safety, and environmental issues, is characterized by the formal and quantitative definition of risk and cost-benefit analyses, where potential damage should be known and measured so that it can either be regulated or banned. In this model, risk is calculated as the product of the probability and severity of its consequence to define how much risk is "acceptable" in practice in exchange for the potential benefits generated[27]. Under these circumstances, according to Boyd (2012), the notion of safety in terms of consumer protection was explicitly redefined as that of "acceptable risk", such as the U.S. Food and Drug Administration's (FDA) definition of 100 million as an "acceptable" amount for a carcinogen called diethylstilbestrol in the 1970s.

Another approach that emerged in the US, the risk regulation model establishing democratic oversight for problems that affect the entire population, has as its main example the National Environmental Policy (NEPA) which requires a risk assessment to be conducted before a project begins, with its subsequent publication for the general population to discuss[28]. Likewise, Brazil also has a rich structure for democratic oversight of problems that affect the entire population, as in the case of environmental issues. In this context, the Brazilian legal system offers different ways for civil society to act in en-

27 KAMISNKI, 2022, p. 36; BOYD, Willian. Genealogies of Risk: Searching for Safety, 1930s–1970s. Ecology Law Quarterly, nº 895, 2012. Disponível em: https://scholar.law.colorado.edu/faculty-articles/143/.

28 KAMINSKI, 2022.

vironmental protection, such as participation in the formulation and implementation of environmental policies, either through the participation of civil society representatives in collegiate bodies with normative powers, or through the possibility of participating in public hearings in the context of environmental impact studies, or even through participation in municipal[29] environmental councils[30].

The third model began in the UK[31] and spread internationally in the 2000s. The regulation focuses on a central administration that assesses risks at a macro level and allocates resources accordingly. In this model, regulators identify the risk to be managed; select their tolerance level; assess the damage and the likelihood of its occurrence; establish risk scores for companies and activities (such as "high," "medium" or "low"); and link the allocation of compliance and inspection resources to these risk[32]. In this respect:

> "Regulators, and the regulatory system as a whole, should use comprehensive risk assessment to concentrate resources on the areas that need them most"
>
> (...)
>
> "Regulators should recognize that a key element of their activity will be to allow, or even encourage, economic progress and only to intervene when there is a clear case for protection."
>
> (HAMPTON, 2005, p. 7, tradução própria)

And lastly, in enterprise risk management, companies self-organize to reduce their risks, such as liabilities or other penalties, whether market or regulatory. To this end, they conduct cyclical and continuous risk assessments, based on an organizational

29 Environmental councils exist at the federal level (National Environmental Council (CONAMA)), in the state level (State Environmental Council (COEMA)) and in the municipalities (Municipal Environmental Council (CONDEMA)) as part of the National Environmental System (Sisnama – created by Law 6.938/81). These councils are normative collegiate bodies – i.e., with robust powers to propose rules and guidelines on environmental management – and are made up of representatives from public bodies, the business sector and civil society; FIGUEIRA, Paulo Sérgio Sampaio. The role of the environmental council in public environmental policies. Published on April 14, 2022. Available at: https://direitoambiental.com/o-papel-do-conselho-do-meio-ambiente-nas-politicas-publicas-ambientais/.

30 COLOMBO, Silvana. The mechanisms of public participation in environmental management in light of the constitutional text: positive and negative aspects. Publisher Unijuí: Human Rights and Democracy Magazine, year 9, no. 18, July/Dec 2021.

31 The English school of regulation studies can be understood as a set of theories, studies and research that have been developed in the United Kingdom at the Centre for Analysis of Risk and Regulation (CARR) da London School of Economics, in which authors such as Christopher Hood and Julia Black are mentioned; ZANATTA, Rafael A. F. Personal Data Protection as Risk Regulation: a new technical framework? FIRST MEETING OF THE INTERNET GOVERNANCE RESEARCH NETWORK, NOVEMBER 2017. Available at: https://www.redegovernanca.net.br/public/conferences/1/anais/ZANATTA,%20Rafael_2017.pdf. p. 182.

32 KAMINSKI, 2022, p. 37.

culture of risk reduction from product design to post-implementation[33]. This model is closer to the notion of self-regulation, because, unlike the other models mentioned, the state is less present when it comes to risk management, since risk management can take place on the institutions' own initiative in the absence of regulation or through state incentives in cases of publication of recommendations, oversight, or the threat of regulatory enforcement[34].

What can be observed is that risk analysis and mitigation can be conducted at a micro level (of the company and a sector, for example) or at a macro level (as in markets in general, in a more cross-cutting manner, with the participation and intervention of the state). At the micro level, the process usually begins with an analysis of the system to identify risks, followed by their mitigation and, at the end, a test to ensure that this mitigation has been effective. This process can be continuous, including analysis of the technology's behavior once it has been introduced into the market. At the market level, risk analysis and mitigation become a regulatory approach. Regulators, public and private entities[35], identify risks and catalog certain companies or activities with different levels of risk. Based on this assessment, the regulator's inspection and investigation resources are deployed to monitor business activity[36].

To fully grasp the international landscape of risk regulation and bring it into the AI context, it is also necessary to look at the risk-based approach adopted by other fields. This is the case with health, food, the environment, insurance, consumer and, more recently, personal data protection, since risk regulation "regimes" can vary from one domain to another even changing over time[37]. Within this context, risk can be understood as a central element that focuses on processes such as the collection of information and cognition of risks; the development of rules and standards of conduct; and the enforcement and monitoring of behavior modification in accordance with the standards created[38].

In the particular case of personal data protection, the European Union's General Data Protection Regulation (GDPR) is an example of risk-based regulation, which provides for a flexible version of a bottom-up regulation (or flexible top-down), responsive to the regulated entities and allocating the regulators' supervisory resources by risk. In

33 Ibid.

34 Ibid, p. 36.

35 According to Julia Black, regulation can be seen from a decentralized perspective, in what she conceptualizes as polycentric or multimodal regulation, i.e. not dependent solely on state forces, but arising from many forums (national, subnational and international), including non-state actors such as companies, organized civil society, people who control the main resources that companies need (for example, credit rating agencies, insurers, auditors, Internet service providers, etc.), "political entrepreneurs", among others; BLACK, Julia. *Proceduralization and polycentric regulation*. Law Magazine GV, Special 1, pp. 099-130, 2005. p. 105-110.

36 KAMINSKI, 2022.

37 Hood *et al*, 2001, p. 8.

38 Ibid.

this sense, Quelle[39], in defining GDPR's risk-based approach, states that it "introduces the notion of risk as a mandatory reference for calibrating the legal obligations of controllers"[40]. In other words, the risk-based approach affects the controllers' obligations in each specific case, which means that data protection law applies differently depending on the level of risk of a given activity[41].

This approach does not replace the principles and rules of data protection with a mere risk analysis. Based on the degree of risk, considering severity and probability, the obligations of each of the controllers are determined with more or less obligations, rights, and duties: the greater the risk, the greater the burden of responsibility[42].

Thus, in the GDPR's risk-based approach, the possible results of certain data processing are highlighted, in order to assess whether the rights and freedoms of individuals are being respected under the terms set by law[43]. This assessment, in high-risk situations, is, for example, embodied in the obligation to draw up a DPIA (Data Protection Impact Assessment), one of the means of impact assessment, considered a key tool in the risk-based approach, which will be further explored in section 2 of this study.

That being said, using the context of data protection as an illustration, the idea behind asymmetric risk regulation is to calibrate the weight of regulation, that is, the intensity of obligations, rights and responsibilities of a given regulated agent, according to the level of calculated risk.

Asymmetric regulation and risk in Brazilian regulatory proposals on AI

In the Brazilian context, the final version of Bill 21/20 approved in September 2021 by the Chamber of Deputies features a fundamentally principled and concise regulation, with only 16 articles. The term "risk" appears in only three of them: (a) Article 2, VI – presents a definition of the artificial intelligence impact report; (b) Article 6, V – in the definition of the safety principle; and (c) Article 9, IV – stipulates that AI agents have a duty to implement artificial intelligence systems after assessing their objectives, benefits and risks related to each phase of the system.

39 QUELLE, Claudia. 'The 'risk revolution' in EU data protection law: We can't have our cake and eat it, too' in R Leenes, R van Brakel, S Gutwirth and P De Hert (eds), Data Protection and Privacy: The Age of Intelligent Machines (Hart Publishing, forthcoming). 2017.

40 "The relationship between the risk-based approach and adherence to the legal requirements of data protection is addressed in particular by articles 24, 25(1) and 35 of the GDPR. These provisions determine how controllers should give hands and feet to data protection law in practice." ("Data Protection and Privacy : The Age of Intelligent Machines", 2017, p. 8); QUELLE, 2017, p. 1.

41 QUELLE, 2017.

42 QUELLE, 2015; ZANATTA, 2017.

43 QUELLE, 2017.

Risk is not used in a systematic way to organize the regulatory approach, but merely as one of several other conceptual elements of the project. Therefore, in this case, there is no need to discuss risk-based regulation, let alone the different degrees of risk of each AI system, since the bill does not proceduralize it by not providing a basic definition of the possible degrees of risk and governance instruments. This conclusion can also be drawn from the analysis of Bill 872/21, which mentions "risks" only once when it defines, in item VII of Article 4, that AI solutions must "follow governance standards that ensure the continuous management and mitigation of potential technology risks", but without providing further explanation of what these processes would consist of. Bill 759/23 makes no mention of the term risk. It contains 7 Articles, divided into: principles (Article 2), guidelines (Article 3), criteria to be met by artificial intelligence systems (Article 4), the obligation of the Executive Branch to create a National Artificial Intelligence Policy (Article 5) and the power of public entities to enter into agreements with private or public, national, or international entities to support and strengthen the National Artificial Intelligence Policy (Article 6). There is no mention of the term "risk" in Bill 5051/19, which is another example of a bill with a principle-based regulatory approach.

Bill 2338/23 combines two different approaches: rights-based and risk-based. The rights-based approach allows for the protection of natural persons impacted by artificial intelligence systems, while the risk-based approach, by regulating the governance of artificial intelligence systems, guarantees predictability and legal certainty for innovation and technological development. This alliance seeks to harmonize the protection of fundamental rights and freedoms, the valorization of work and human dignity with the creation of new value chains and the development of the economic order:

> In terms of structure, the proposal establishes risk-based regulation and rights-based regulatory models. It also outlines governance instruments for proper accountability of economic agents who develop and use artificial intelligence, encouraging good faith and effective risk management.
>
> (Bill 2338/23, p.29)

The model adopted by Bill 2338/23 uses risk, like most recent AI regulation initiatives (see: table below), to standardize the responsibility and obligations of the AI system's agents. There are basic rights that apply to any interaction between the AI system and a human being (as per Article 5, I, II, IV, V and VI, Article 7, Article 8, Article 12) in the spirit of a rights-based regulation, such as information and transparency. However, there are more obligations when there is a greater risk to rights (Article 5, III; Article 9;

Article 10; Article 11). Similarly, governance measures for artificial intelligence systems are also divided according to risk:

> Beyond establishing basic and cross-cutting rights for any context in which there is interaction between machine and human being, such as information and transparency, this obligation is intensified when the AI system produces significant legal effects or has a significant impact on the subjects (e.g., the right to contest and human intervention). As such, the weight of regulation is adjusted according to the potential risks in the context of the technology's application. Certain general and specific governance measures have been established for artificial intelligence systems with any degree of risk and those categorized as high risk, respectively.
>
> (Bill 2338/23, p.30-31)

Separation of governance measures according to the risk of a given AI system in a specific context resembles the risk-based regulation originated in the UK by the English school of regulation[44], described by Kaminski, and the approach of the EU AI Act. In this version of risk regulation, historically, there has been provision for state oversight that co-assesses and co-assigns risks to certain companies and activities, dividing them, for example, between high, medium, and low, and allocates law enforcement and investigation resources according to these risks[45]. For example, Bill 2338/23 provides that, in addition to the necessary designation of a competent authority (Article 32, head provision), this regulatory body must cooperate with others with related competencies in order to understand and manage risks (Article 32, items V, VII and VIII). In particular, when it comes to specific economic sectors in which there will be contextual variation of risks in AI development and implementation (Article 34, head provision and §1).

Furthermore, Bill 2338/23 differentiates itself qualitatively from other national regulatory proposals by providing for a chapter on governance and good practices in order to encourage economic agents themselves to manage the risks of their own

44 The English school of regulation studies can be understood as a set of theories, studies and research that have been developed in the United Kingdom at the Centre for Analysis of Risk and Regulation (CARR) da London School of Economics, in which authors such as Christopher Hood and Julia Black are mentioned; ZANATTA, Rafael A. F. Personal Data Protection as Risk Regulation: a new technical framework? FIRST MEETING OF THE INTERNET GOVERNANCE RESEARCH NETWORK, NOVEMBER 2017. Available at: https://www.redegovernanca.net.br/public/conferences/1/anais/ZANATTA,%20Rafael_2017.pdf. p. 182.

45 KAMINSKI, 2022.

economic activities. It also expressly states that it is up to the regulatory authorities to promote "studies on good practices in the development and use of artificial intelligence systems" and "experimental regulatory environments" (sandboxes). This public-private partnership arrangement in understanding the risks associated with AI is further complemented by the state apparatus' duty to be accountable for its regulatory choices (e.g. regulatory impact assessments and public consultations) and, as will be shown below, by the most participatory model possible for drawing up algorithmic impact assessments, which will be one of the documents making up a public and open database on high-risk AIs.

In short, as the table below summarizes, Bill 2338/23 not only quantitatively lists risk as an organizing element of regulation. It also qualitatively proceduralizes the way in which institutional resources and tools should be allocated to democratically decide which risks are (un)acceptable and how to manage them.

RISK REGULATION MODELS ACCORDING TO THE AI REGULATION PROPOSAL OR INTERNATIONAL DOCUMENTS[46]

Explanatory memorandum to the proposal / *Soft law* document explained		Quantitative model	Quantitative Model (Potentially Hybrid)		
			Democratic oversight	Expected allocation of resources (e.g., creation or designation of regulatory body)	Corporate risk regulation
Bill 21/20	Bill 21/20 by Congressman Eduardo Bismarck mentions the term 'risk' when referring to duties related to the management of risks caused by AI systems, in combination with their potential benefits. The regulation seeks to "make internationally recognized principles mandatory and regulate rights and duties," encouraging the use of AI to "promote research and innovation, increase productivity, develop sustainable economic activity, improve people's well-being and help respond to the main global challenges." However, the text approved in the Chamber of Deputies comes from the substitutive by representative Luísa Canziani which, despite amending the bill, has no explanatory statement.	Low-Nonexistent	Low-Nonexistent	Low-Nonexistent (It only suggests guidelines for action by the public authorities and coordination of existing sectoral authorities. - e.g. Article 6)	Average (e.g, Articles. 3, VIII, and 8, III)

46 This table does not provide a conclusive, opinionated comparison of all the regulatory proposals, but it does help with the initial mapping of their relevant points in order to move forward in this type of comparison, where a proposal can be more or less hybrid according to the risk regulation modalities classified by Kaminski (2022).

Bill 2338/23	CJSUBIA's explanatory statement in the report presented in December 2022 and the grounds for Bill 2338/23 point out that "the proposal establishes a risk-based regulation and a regulatory model based on rights. The proposal seeks to "reconcile, in terms of legal discipline, the protection of fundamental rights and freedoms, the importance of work and the dignity of the human person and the technological innovation represented by artificial intelligence." What this means is that the weight of regulation is dynamically "gauged according to the potential risks of the context in which AI is applied." To this end, "certain general and specific governance measures have been established, symmetrically to the rights, for artificial intelligence systems with any degree of risk and for those categorized as high risk, respectively." In other words, the regulatory burden (greater number of legal obligations) increases as the level of risk of the AI system increases.	**High** (Provides for a broad taxonomy of risks: see table in topic a.2 below - Degrees of Risk, Excessive Risks and High Risk - e.g., Articles 14 and 17)	**High** (Provides for a series of obligations of social control and public participation in regulatory production and risk management - e.g., Chapter IV on governance of AI systems)	**Average** (In addition to providing Guidelines for action by the Public Authorities and coordinating existing Sectoral Authorities, there is provision for a new authority to coordinate such efforts on behalf of the Executive Branch - e.g., Article 21 and Section I of Chapter VIII)	**High** (e.g, Chapter VI - Codes of Good Practice and Section Dedicated to Fostering Innovation)
EU AI Act	The reasons and objectives of the Proposal emphasize that the European Union "is committed to achieving a balanced approach." In this sense, they emphasize that "it is in the EU's interest to preserve the technological leadership of the EU and to ensure that new technologies, developed and exploited with respect for the values, fundamental rights and principles of the EU, are at the service of European citizens."	**High** (Provides for a broad taxonomy of risks: see table in topic a.2 below – Unacceptable and High - e.g., Articles 5 and 6 + Annex III)	**High** (e.g, Recital 81 and Article 29a (4) - EP text version)	**High** (e.g, different governance and implementation measures in Titles VI, VII and VIII, respectively, with the creation of the European Artificial Intelligence Council, coordination between this Council and other national	**High** (e.g, Title IX on Codes of Conduct; Title V on measures to support innovation)

EU AI Act	There is reference to the adoption of a "well-defined risk-based regulatory approach that does not create unnecessary restrictions, since legal intervention is adapted to concrete situations in which there is a justified cause for concern or when such concern can reasonably be anticipated in the near future", as well as bringing in "flexible mechanisms" for dynamically adapting regulation in line with technological advances and the emergence of new concerning situations.			authorities and post-market monitoring measures).	
Directive on Automated Decision-Making + Algorithmic Impact Assessment tool (Canadá)	The Directive explains that its aim is to ensure that automated decision-making systems are implemented in such a way as to reduce the risks to Canadian society. To this end, it provides for the obligation to conduct an algorithmic impact assessment, including providing a practical tool.	**High** (Provides for a broad taxonomy of risks: see table in topic a.2 below - e.g., Annex B)	**High** (mentions the need for consultation with internal and external stakeholders on the tool page)	**Average** (designates existing authority - e.g., Article 2)	**Low-Nonexistent** (in principle, it only applies to automated decisions used for administrative decision-making - e.g., Article 5)
AIDA	AIDA's explanatory document mentions that the regulation will take a "risk-based" approach, in order to align with other regulations under development at international level. The aim is to build a "framework to ensure the initiative-taking identification and mitigation of risks in order to prevent harm and discriminatory outcomes, while recognizing the unique nature of the AI ecosystem and ensuring that responsible research and innovation are supported." To this end, the document goes on to define that "as technology evolves, new AI capabilities and uses	**High** (explicit mention of high-risk AI systems and prohibited practices)	**High** (AIDA's text has not yet been released, but its study document clearly mentions the "extensive consultation with a range of stakeholders" for the construction of the regulation)	**High** (they mention that there are authorities already, but that the risks of AI create the need for new actions + designation of the Department of Innovation, Science, and Industry as the competent authority to implement and supervise AIDA)	**High** (there is evident concern to create proportional regulation that does not hinder innovation)

AIDA	will emerge and Canada needs an approach that adapts to this ever-changing landscape." Furthermore, the explanatory document of the AIDA's proposal stresses that the mandatory measures for AI actors will be determined according to the "context and risks associated with specific regulated activities within the life cycle of a high-risk AI system." Accordingly, "the regulated activities defined in AIDA would be associated with distinct obligations that are proportionate to the risk," avoiding "undue impacts on innovation."				
Algorithmic Accountability Act EUA	O projeto visa instruir a Federal Trade Commission a exigir avaliações de impacto de sistemas de decisão automatizados e processos de decisão críticos. Assim, pretende exigir que as empresas avaliem os impactos dos sistemas automatizados que utilizam e vendem, além de criar uma nova transparência sobre quando e como os sistemas automatizados são utilizados e capacitar os consumidores a fazerem escolhas informadas sobre a automatização de decisões críticas.	**Average** (does not have a well-defined taxonomy of risks, but directs regulation towards higher risks)	**High** (when conducting impact assessments, important stakeholders should be consulted - e.g., Section 3, (b) 1. (G); Section 4, (a) (2))	**High** (directs the Federal Trade Commission (FTC)) to require impact assessments of agents using automated decision systems, in addition to creating obligations for other authorities - e.g., Sections 8 and 9)	**Average** (possibility of technical assistance and guidance from actors regulated by the FTC - e.g., Section 7)
Bill 15869/19 (Chile)	The project is inspired by the EU AI Act and therefore also presents a risk-based approach, grounded on the need to tackle the rapid advance of technologies, both in their positive aspects and in the risks associated with their use. The goal is to "establish an area of digital sovereignty for artificial intelligence systems, in	**High** (Provides for taxonomy of risks: see table in topic a.2 below - Degrees of Risk, Unacceptable Risks and High risk - e.g.	**High** (Provides for a series of obligations of social control and public participation in regulatory production and risk management - e.g., Chapter IV on	**Average** (In addition to setting out Guidelines for action by the Government and coordinating existing Sectoral Authorities, there is provision for a new authority to coordinate such efforts on behalf of the	**High** (e.g., Chapter VI - Codes of Good Practice and a Section Dedicated to Fostering Innovation)

Bill 15869/19 (Chile)	which the State of Chile discusses ethical and legal considerations, as well as regulating the risks arising from the development, distribution, commercialization and use of this technology" and to establish limits, formalities and implementation and application requirements for anyone conducting their actions with the technology. To achieve this goal, the bill establishes the creation of the National Artificial Intelligence Commission, which will have, among its competencies, to propose the expansion or updating of AI regulations; to evaluate and authorize (or prohibit) AI systems; and to keep a register of authorized systems.	Articles 3 and 4)	governance of AI systems)	Executive Branch - e.g., Article 21 and Section I of Chapter VIII)	
Draft [Framework] Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (CAI)	The draft text of the Convention developed by the Council of Europe's Artificial Intelligence Committee (CAI) includes a specific article on the "risk-based approach", in which it defines that each Member State "shall maintain and take graduated and differentiated measures in its domestic legal framework, as necessary and appropriate, in view of the severity and likelihood of adverse impacts on human rights and fundamental freedoms, democracy and the rule of law during the design, development, use and discontinuation of artificial intelligence systems".	**Average** (Directs the use of a risk-based approach, although it does not define these levels - e.g., Article XX)	**High** (provision for diverse and adequate public discussion and consultation - e.g., Article 19)	**Average** (e.g., Chapter VII on follow-up mechanisms and cooperation in the implementation of the Convention)	**Average** (e.g., Article 12 on safe innovation)

	In addition, the bill stipulates that Member States "must take measures to identify, assess, prevent and mitigate risks and impacts on human rights, democracy and the rule of law arising from the design, development, use or discontinuation of AI systems," taking into account the risk-based approach.				
CAHAI	In the regulatory feasibility study, CAHAI explains that the risks posed by AI systems depend on the context of the application, the technology and the stakeholders involved. Therefore, to combat any stifling of AI innovation and to ensure that the benefits of this technology can be reaped while adequately addressing its risks, CAHAI recommends that a future legal framework created by the Council of Europe on AI should follow a risk-based approach. In addition, the Committee also stresses that, where relevant, a preventative approach should be considered, including possible bans. As such, according to the study, a comprehensive legal framework for AI systems, guided by a risk-based approach, can help provide the contours in which beneficial innovation can be stimulated and enhanced, and the benefits of AI can be optimized, ensuring - and maximizing - the protection of human rights, democracy, and the rule of law through effective legal remedies.	**High** (feasibility study mentions the need for a risk-based approach with the definition of degrees of risk and possible prohibitions)	**High** (involvement of stakeholders in the preparation of AI impact assessments)	**High** (need for national AI authorities)	**Average** (e.g., mentions compliance measures such as sandboxes, but combined with impact assessments)

Blueprint for an AI Bill of Rights	The explanatory section on the application of the Blueprint stresses that the measures taken to implement the vision presented in the document must be proportionate to the extent and nature of the damage, or risk of damage, to rights, opportunities and access resulting from AI systems.	**High** (there is no well-defined taxonomy, but it provides for different recommendations based on the risk of the AI system)	**High** (e.g., Principle of Safe and Effective Systems mentions the need for consultation with different actors)	**Average** (the document is at first voluntary - but becomes mandatory for federal government agencies after President Biden's Executive Order of October 30th, 2023)	**High** (the document is intended to be voluntary for organizations, in addition to providing support measures for innovation)
NIST – AI RMF 1.0	The framework's Executive Summary mentions that AI risk management is an essential component for the responsible development and use of the technology. The central objective of the NIST AI RMF is to be an auxiliary source of risk management for organizations that develop, implement, and use AI. The tool is, at the outset, voluntary, rights-preserving, non-sector specific and "use case agnostic," providing flexibility for organizations of all sizes and from all sectors, including allowing for adaptation throughout the development of the technology.	**Average** (defines that the AI RMF can be used to prioritize risk, but does not define risk tolerance, which must be defined by each organization according to how much risk it is willing to assume, but mentions the possibility of risks: unacceptable, high, and low - topic 1.2.2)	**High** (e.g., topic 5.2, page 25 - mentions a requirement for the incorporation of diverse internal teams and the involvement of different external agents, including individuals and groups potentially impacted by the technology.)	**Low-Nonexistent** (this is a voluntary document, initially)	**High** (the framework is intended to be applied voluntarily in different organizations of various sizes and sectors)
OCDE	The OECD has been following the development of AI governance since 2019, with the publication of its **principles** and the creation of the **AI Observatory**, with a strong concern for the economic and social impacts of this technology. It has produced various reports and studies on the regulation of AI, in particular the **framework for the classification of**	**High** (has already come out in favor of a risk-based approach to regulating AI, in order to direct oversight and intervention	**High** (need for monitoring and stakeholder participation in AI accountability processes)	**High** (need for supervising authorities to monitor AI policies)	**High** (different principles and good practices to be implemented by organizations for a reliable AI system)

OCDE	AI systems and the report on accountability measures, as well as the recent guideposts to promote interoperability in AI risk management, which demonstrate the international trend towards regulating AI systems using a risk-based approach with the need for accountability tools.	where they are most needed, while avoiding unnecessary obstacles to innovation)			
UNESCO	UNESCO produced the first-ever global standard on AI ethics – the "Recommendation on the Ethics of Artificial Intelligence," published in November 2021, based on a human rights approach. This framework was adopted by all 193 Member States. The protection of human rights and dignity is the cornerstone of the Recommendation, based on the advancement of fundamental principles such as transparency and fairness, always remembering the importance of human oversight of AI systems. The document is based on four core values: promotion of human rights and human dignity; Living in peaceful, just, and interconnected societies; Ensuring diversity and inclusiveness; and Environment and ecosystem flourishing. In addition, it follows recent initiatives emphasizing the need to go beyond ethical principles to effective practical strategies. To this end, the Recommendation creates 11 key areas for policy action ("actionable policies") and provides two practical methodologies for (i) ethical impact assessment (EIA); (ii) readiness assessment.	**High** (e.g., Articles 25 and 50-53 of the Recommendation + Ethical Impact Assessment Tool brings 4 levels of risk: extremely high, high, average, and moderate/low)	**High** (e.g., Articles 50-53 of the Recommendation)	**High** (e.g., Chapter V of the Recommendation on monitoring and evaluation)	**Average** (e.g. Article 69)

Based on these comparisons, the alignment of Brazil's Bill 2338/23 with international discussions stands out, not only from the European context, but also from global standards, such as the OECD and UNESCO. This approach highlights the risk regulation model defined by Kaminski (2022) in the form of democratic oversight and prior allocation of resources, since most AI governance proposals focus on the need for public participation in regulatory processes. That is especially true when it involves conducting impact assessments, as well as paying great attention to the division of regulatory efforts according to the systems' risks. In addition, there is provision for regulatory tools to unlock a kind of public-private partnership in risk management, a middle ground between regulation monopolized by the state (command and control) or solely by the economic agent itself (self-regulation)[47].

a.2) Risk Taxonomy

The risk regulation model, or asymmetric regulation, intensifies both the resources used by the regulator for oversight and the obligations that companies must fulfill in relation to the products or services that present the greatest risk. In this model, risks are allocated in a macro way to certain companies or activities, creating risk bands which can vary according to the methodology chosen, being divided into "high," "average" or "low."

The same strategy of regulation and macro designation of risk bands has been adopted by various regulations governing artificial intelligence systems: Bill 2338/23 in Brazil, the EU AI Act, Chile's *Proyecto de ley* 15869/19[48] and Canada's Algorithmic Impact Assessment tool under Canada's Automated Decision Making Directive.

In Brazil, specifically in Bill 2338/23, risk is divided into three bands: excessive, high, and moderate/low (this is a residual category, which is not explained in the legislation). There is no definition of each category of risk, since it is classified according to a list of examples, with the provision of quantitative and qualitative elements for updating the list of unacceptable and high risk systems by the competent authority, according to Article 18.

In the European Union, the proposed regulation (EU AI Act) categorizes risk along

47 For the purposes of analyzing the relationship between democratic supervision and regulatory models, see, among others: a) how the precautionary principle is organizational for the purposes of public deliberation on what the acceptable risks of an economic activity or technology are. (BIONI, Bruno; LUCIANO, Maria. The Precautionary Principle for the Regulation of Artificial Intelligence: Would Data Protection Laws be its Gateway? In: Frazão, Ana. Mullholand, Caitlin. Artificial Intelligence and Law: ethics, regulation, and responsibility. São Paulo: Revista dos Tribunais, 2019); b) the notion of informational co-deliberation in addition to that of informational self-determination (BIONI, Bruno Ricardo. *Regulation and Personal Data Protection – The Principle of Accountability*. São Paulo: Editora Forense, 2022. 320p).

48 https://www.camara.cl/verDOC.aspx?prmID=72777&prmTipo=FICHAPARLAMENTARIA&prmFICHATIPO=-DIP&prmLOCAL=0.

the same lines as Bill 2338/23, in three categories, although the highest level of risk is called unacceptable. The lowest risk band (moderate or limited), as in Bill 2338/23, is not spelled out in the legislation and is a residual category. This classification is also followed by the Chilean Bill 15869-19, which separates risk into unacceptable and high, in addition to the residual category of systems not classified by the two levels of risk.

Canada's Algorithmic Impact Assessment tool, within the framework of the Canadian Directive on Automated Decision Making, splits the impacts of automating an administrative decision into 4 levels, with each level having a percentage range of impact, depending on the reversibility of the automated decisions and the expected duration of the decision taken. Therefore, reversible, and brief decisions have little impact (level I), and irreversible and perpetual decisions have an extremely high impact (level IV).

Regulation	Risk Levels	Nomenclature	Obligations
Bill 2338/23	3	Excessive	Prohibition (Article 14)
		High	Documentation, conducting reliability tests, adopting technical measures to ensure that the results can be explained, among others (Article 20) + general obligations (Article 19) + obligation to draw up an algorithmic impact assessment (Article 22)
		Moderate / low (residual category)	Transparency measures, adequate data management measures to mitigate and prevent discriminatory bias, information security measures by design, among others (Article 19).
EU AI Act	3[49]	Unacceptable	Prohibition (Article 5)
		High	Obligation to draw up an impact assessment on fundamental rights (Article 29a), quality management system, drawing up technical documentation, keeping records, being subject to the conformity assessment procedure, adopting corrective measures, among others (chapter 2 and 3 - articles 8 onwards), among others.
		Moderate	Rather limited transparency obligations, for example with regard to the provision of information to signal the use of an AI system when it interacts with humans (Title IV).

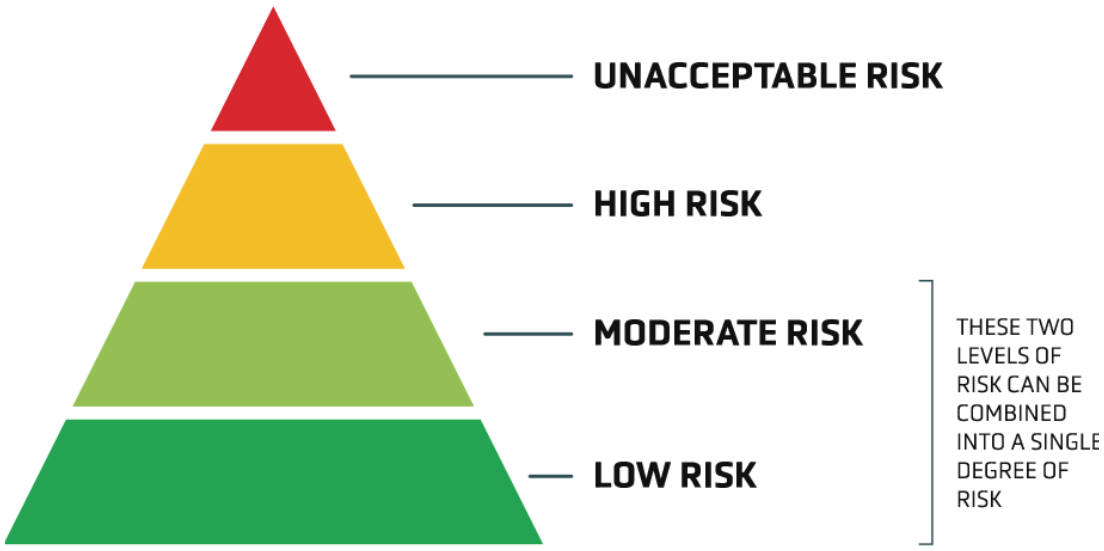
49 The latest version of the text, coming from the European Parliament in June 2023, created specific obligations for providers of foundation AI models in Article 28b, in addition to the issue of associated risk.

Bill 15869/19 (Chile)	3	Unacceptable	Prohibition through non-authorization by the competent authority (Article 8).
		High	Prior obligations, such as the implementation of a risk management plan, input data management plan and quality management plan, record keeping, providing information, human intervention, among others (Article 9).
		Residual	Inform people that they are interacting with an AI system (Article 10) and inform the authority in the event of a serious incident or malfunction (Article 11).
Directive on Automated Decision-Making + Algorithmic Impact Assessment tool	4	Low to no impact (level I)	Substantial explanation for the results of automated decisions.
		Moderate impact (level II)	*Peer review* (consultation with at least two experts and disclosure of the summary results on the Canadian Government's website); gender analysis, plain language notice published on all service delivery channels in use; relevant explanation provided to the client in any decision that results in the denial of a benefit or service; documentation of the design and operation of the system, among others.
		Impacto Alto (level III)	*Peer review* (consultation with at least two experts and disclosure of the summary results on the Canadian Government's website); decisions cannot be made without human intervention; gender analysis, plain language notice published on all service delivery channels in use; meaningful explanation provided to the customer in any decision that results in the denial of a benefit or service; documentation of the design and operation of the system, operation depends on Deputy Head approval, among others.
		Impacto muito alto (level IV)	*Peer review* (consultation with at least two experts and disclosure of the summary results on the Canadian Government's website); decisions cannot be made without human intervention; gender analysis, plain language notice posted on all service delivery channels in use; meaningful explanation provided to the customer in any decision that results in the denial of a benefit or service; documentation of system design and operation, recurrent training courses (and a means to verify that training has been completed); operation depends on Treasury Board approval, among others.

Therefore, different international regulatory initiatives from Europe, Canada and Latin America apply the risk-based approach, dividing it into degrees, levels, or ranges so that regulatory resources, as well as the burden of obligations on regulated agents, are properly distributed among them. In the Brazilian context, this is only defined in Bill 2338/23, which breaks down the degrees of risk for the allocation of legal obligations, as well as indicated where regulation should be more intense, something which isn't addressed in the other Brazilian bills currently in dispute in the country.

a.3) Risk levels

As previously mentioned, through the taxonomy of risks in asymmetric risk regulation, there are degrees of risk (different levels of risk measurement), which can vary according to the benchmark used. The different degrees of risk will give rise to stronger or weaker regulatory obligations. The idea is not to create unnecessary restrictions on trade, the provision of services or innovation by means of legal intervention tailored to concrete situations in which there is a justified cause for concern or in which such concern can reasonably be anticipated in the near future[50].



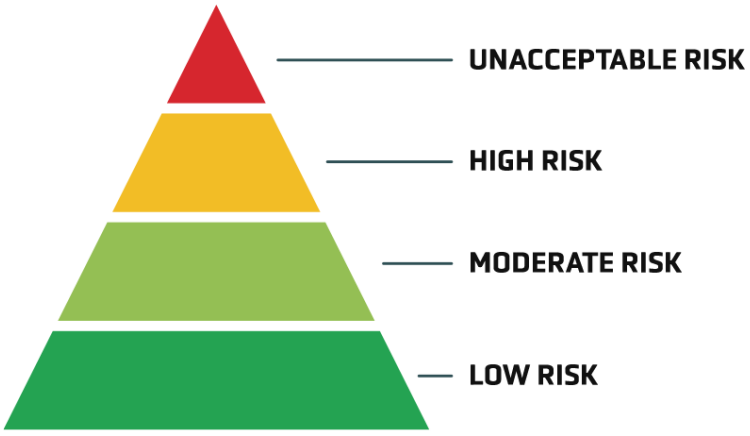
Risk gradations (different levels of risk measurement) can vary according to the benchmark used.

Thus, the gradation of risks will unlock different regulatory weights according to the risk identified for those developing the technology and launching it on the market

50 EU AI Act, explanatory memorandum.

or putting it to use. Regulatory measures will increase in intensity the greater the risk associated with their launch. At this point, it is possible to relate the gradation of risks to the idea of a pyramid, where the base is the minimum/average risk (without major responsibilities), the middle is the high risk (there are many requirements so that the technology is allowed to be implemented on the market) and, finally, the top where there are unacceptable risks, i.e. where regulation will prevent the use of the technology because it brings more risks than benefits for rights, democracy and society.

However, in addition to defining the different levels of risk, it is essential to provide qualitative elements for defining each of these risks. In other words, instead of just defining the degrees of risk (e.g. low/average/high risk) generically, it is essential to have minimum criteria for identifying the systems at each of these levels. By way of example:



UNACCEPTABLE RISK
When there are "significant" or unknown risks to human rights, democracy, or the rule of law + there are no appropriate mitigation measures from existing legal frameworks to adequately mitigate.

HIGH RISK
When there are "significant" or unknown risks to human rights, democracy, or the rule of law + there are appropriate mitigation measures within existing legal frameworks that can assist in adequate mitigation.

MODERATE RISK
When, although there are risks, they are not significant for human rights, democracy, or the rule of law + there are appropriate mitigation measures within existing legal frameworks that assist in mitigation.

LOW RISK
The presence of zero risk is impossible, but existing risks are not significant for human rights, democracy, or the rule of law, and are only residual, not requiring additional requirements for their use.

Risk gradations (different levels of risk measurement) can vary according to the benchmark used.

For example, it is essential to consider the context in which AI technology is applied, since it is from this application context that a more granular analysis for parameterizing risk is possible. In addition to context, other elements can be used as criteria for defining the degree of risk, such as scope, explainability, amount of data processed, level of automation, among others, which should be minimally explained in the regulation.

a.3.1) *Unacceptable/Excessive Risk*

According to Mantelero (2022), whenever a new application of technology can produce serious potential risks for individuals and society, which cannot be calculated or quantified precisely and in advance, a precautionary approach should be adopted. In these cases, the implementation of governance mechanisms, such as drawing up a

proper impact assessment, is impossible, but the potentially high impact on society justifies specific precautionary measures (for example, a ban or restriction on the use of the technology). It is at this level of risk that the most interventionist and regulatory level is found since the ban on use and development of systems is defined *ex-ante*.

Artificial intelligence systems of unacceptable and excessive risk prompt stronger regulatory interventions from the EU AI Act, Chilean Bill 15869/19, and Bill 2338/23, respectively, resulting in the prohibition of these systems in advance (*ex-ante*). This level can be understood as part of the reconciliation between a rights-based and risks-based approach. In other words, there are certain rights that are non-negotiable and certain applications of artificial intelligence would generate intolerable risks.

An example of this prohibition are social scoring systems, which condition access to goods, services, and public policies upon an assessment of the individual based on their social behavior or personality traits.

The idea of social credit systems has been disseminated worldwide following experiences in China. In 2014, the Chinese Central Government announced a six-year plan to establish a "social credit system", in which actions that create trust in society would be rewarded, while those that run counter would be punished[51]. The term social credit encompasses not only what is traditionally seen as a credit score, i.e., the financial history of individuals and companies and a prediction of whether they will repay future loans, but also the social creditworthiness score, which relates to an individual's trustworthiness from non-financial activities[52]. At a national scale, what exists for the time being is a system focused on companies, which aggregates data on compliance with regulations from different government agencies, made available on a website called "Credit China"[53]. Despite the focus on companies, there is information on individuals and other organizations on the site, bringing together varied but unsystematized databases with information such as which individuals have failed to comply with judicial measures, which Chinese universities are legitimate, and so on[54].

The most developed examples of social credit systems come from local governments that have implemented pilot programs[55]. In the city of Rongcheng, with half a million inhabitants, a system was implemented in 2013 that gave each citizen 1,000 points as a basis for social credit, where the number of points was influenced by individ-

51 YANG, Zeyi. China just announced a new social credit law. Here's what it means. MIT Technology Review, publicado em 22 nov. 2022. Disponível em: https://www.technologyreview.com/2022/11/22/1063605/china-announced-a--new-social-credit-law-what-does-it-mean/.

52 Ibid.

53 Ibid.

54 Ibid.

55 Ibid.

ual actions, such as spreading malicious information on social media networks, which reduced the number by 50 points, or winning a nationwide sports or cultural competition, which added 40 points[56]. These programs remained restricted to cities, not reaching entire provinces or the country[57]. In fact, in December 2020, in a guide published by the Chinese State Council, it was recommended that local governments only punish behavior that is already illegal under Chinese law. Returning to the example of Rongcheng City, the social credit regulations have been updated to allow citizens to leave the program if they wish and some criteria have been modified[58].

Another relevant example is the use of remote biometric identification systems in public spaces, such as facial recognition in public security. Different studies[59] have already shown that, at the current stage of development, these systems present inaccuracies of false positives and negatives, especially against already marginalized and vulnerable groups, especially when analyzed through an intersectional perspective. Therefore, the implementation and use of this technology for public security purposes by the state, especially when applied massively in real time for identification and tracking, can negatively interfere with different fundamental rights, including reinforcing structural discrimination.

Given this context, there is discussion regarding the need for a ban or moratorium on the development and use of facial recognition systems in public security by the state. In the case of a ban, a total prohibition on the use of these systems is advocated, on the grounds that their benefits do not outweigh the harm caused by the violation of non-negotiable rights and values, such as non-discrimination. In the case of a moratorium, there is a ban for a certain period or under certain circumstances, until technology evolves, or efficient governance mechanisms are developed so that non-negotiable rights are not violated by those systems.

Bill 2338/23, in the excessive risk section, includes as one of its hypotheses the use of remote biometric identification systems in public security activities.

In effect, what the article creates is a moratorium, making the use of such systems conditional on two factors: (i) the enactment of a specific federal law, (ii) judicial authori-

56 Ibid.

57 Ibid.

58 Ibid.

59 Buolamwini, Joy; Gebru, Timnit. (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Conference on Fairness, Accountability and Transparency. Proceedings of Machine Learning Research 81:1–15, 2018. Available at: <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>; COSTANZA-CHOCK, Sasha. Design Justice, A.I., and escape from the matrix of domination. Cambridge: Journal of Design and Science, Jul. 2018. DOI:10.21428/96c8d426. Available at: https://jods.mitpress.mit. edu/pub/costanza-chock/release/4; VARON, Joana; SILVA, Mariah Rafaela. Facial recognition in the public sector and trans identities: techno politics of control and threat to gender diversity in its intersectionality's of race, class, and territory. Available at: <https://codingrights.org/docs/rec-facial-id-trans.pdf>.

zation for use, which must be connected to individualized criminal prosecution activity, for crimes that carry a maximum penalty of imprisonment of more than two years, the search for victims or missing persons, or in case of a warrantless arrest (when a person is caught *"red-handed"*)[60]. Federal law must provide for measures that are proportionate and strictly necessary to serve the public interest, in addition to the need for a review by the public agent responsible for the algorithmic inference before action is taken on the identified person.

The Chilean bill also includes a moratorium on the use of remote biometric identification systems in publicly accessible spaces in cases where this use is considered strictly necessary for: (i) searching for possible specific victims of a crime, including missing minors; (ii) preventing a specific, significant and imminent threat to people's lives or physical safety or a terrorist attack; (iii) detection, location, identification or prosecution of a person who has committed, or is suspected of having committed, any of the crimes provided for in the Penal Code. In these three exceptional cases, the bill also determines that they are always subject to a prior decision by a Court of Justice and only applied by the *Carabineros* of Chile (a kind of ostensible police force) and the Investigative Police.

In European regulations, there are two different stances on biometric identification regulation. The European Parliament's standpoint prohibits the use of real-time remote biometric identification systems in public spaces (by public or private entities), as well as the use of systems for analyzing recordings of public spaces with remote biometric identification. There is an exception for retroactive use, i.e., of recordings for remote biometric identification if there is prior judicial authorization for the use, which must take place in the context of criminal prosecution when strictly necessary and be related to a serious crime that has already taken place[61].

The Council of the European Union, made up of Member States, has weakened the ban on the real-time use of remote biometric identification systems, as it supports the possibility of using them in exceptional situations, listed exhaustively, in which the public interest will prevail over the risks. Examples of such situations include the search for potential victims of a crime, including missing children; some threats to the lives or physical safety of natural persons or a terrorist attack; and the detection, location, iden-

60 Based on the wording of the proposed law, Brazil would prohibit the practice of Predictive Policing, i.e. the use of algorithms to analyze large databases to predict information related to crimes, such as when and where they will happen in the future or who is most likely to commit them, and, based on this information, make a decision on where to allocate a larger police contingent. The reason for this is that the use of remote biometric identification systems in public security activities could only exist in the cases listed in the sections of Article 15, no longer analyzing the general public as suspects, since analysis would only be possible in the context of individualized criminal prosecution. ACLU of Washington. How Automated Decision Systems are used in Policing. Published on 26 Dec. 2022. Available at: https://www.aclu-wa.org/story/how-automated-decision-systems-are-used-policing.

61 Amendment 41 – https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf.

tification, or prosecution of suspects of the 32 crimes listed in the Council Framework Decision 2002/584/JHA if the crimes are punishable in the Member States by a custodial sentence or detention for a maximum period of at least three years. In addition, in the Council's proposal, the use of such systems by border, immigration or asylum police forces to identify a person who refuses to be identified or cannot prove their identity is also permitted[62].

Some international reference documents do not clearly specify which cases of AI pose an excessive risk, but they do provide that, under certain conditions, some AI systems should be subject to a prior moratorium or ban. In the draft Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law of the Committee on Artificial Intelligence (CAI) of the Council of Europe, an obligation is created for States Parties to take the necessary legislative measures to create a moratorium or ban on certain AI systems whenever they are considered incompatible with respect for human rights, the functioning of democracy and the rule of law (Article 15(3)). Meanwhile, UNESCO's Recommendation on the Ethics of Artificial Intelligence mentions the prohibition of AI systems that have disproportionate negative effects on environmental impacts (Article 86), as well as those that have the power to make life and death decisions (Article 36), in addition to the clear mention of the recommendation not to use AI for the purposes of mass surveillance and social credit (Article 26).

62 Paragraph 19 et seq. Available at: https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf.

Regulation	*Ex-ante* prediction of prohibited risks	Number of prohibited situations *ex-post*	Use of biometric data in AI for the purpose of criminal prosecution
Bill 2338/23	Yes	4	Moratorium
EU AI Act (European Parliament version)	Yes	8	Ban
EU AI Act (Council of the European Union's version)	Yes	4	Moratorium
Bill 15869/19 (Chile)	Yes	4	Moratorium
CAI	Yes	There is no definition of quantity	No specific mention
UNESCO	Yes	There is no definition of quantity	No specific mention

ANALYZED REGULATIONS Excessive/unacceptable risk hypotheses				
	Bill 2338/2023	EU AI ACT (Council)	EU AI ACT (EP)	Bill 15869-19 (Chile)
Definition of risk/ consequence (prohibition)	Article 14: The use and implementation of artificial intelligence systems is prohibited when:	Article 5 (1) The following artificial intelligence practices shall be prohibited:	Article 5 (1) The following artificial intelligence practices shall be prohibited:	Article 3. The following shall be classified as AI systems presenting an unacceptable risk: Article 8. The Commission shall not authorize the development, distribution, marketing, or use of AI-systems whose risk is unacceptable
Techniques for inducing behavior that can cause harm (physical and psychological)	I – employing subliminal techniques that have the purpose or effect of inducing a natural person to behave in a manner that is harmful or dangerous to their health or safety or against the foundations of this Law;	a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's behavior in such a way as to cause or be reasonably likely to cause physical or psychological harm to that person or to another person;	a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's or a group of persons' behavior by appreciably impairing the person's ability to make an informed decision, thereby causing the person to take a decision that that person would not have otherwise taken in a manner that causes or is likely to cause that person, another person or group of persons significant harm; The prohibition of AI-systems using subliminal techniques	1. AI-systems that use subliminal techniques that transcend a person's consciousness to substantially alter their behavior in a manner that causes or could cause physical or mental harm to that person or another;.

			referred to in the first subparagraph shall not apply to AI-systems intended for therapeutic purposes approved on the basis of the specific informed consent of the persons exposed to them or, where appropriate, of their legal guardian;	
Techniques that exploit vulnerabilities	II – that exploit any vulnerabilities of specific groups of natural persons, such as those associated with their age or physical or mental disability, in order to induce them to behave in a manner that is harmful to their health or safety or against the foundations of this Law;	b) The placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, disability or a specific social or economic situation, with the objective to or the effect of materially distorting the behavior of a person pertaining to that group in a manner that causes or is reasonably likely to cause that person or another person significant physical or psychological harm;	b) The placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a person or of a specific group of persons including characteristics of the known or predicted personality traits of that person or of a group or social characteristics or economic situation, age, physical or mental capacity with the purpose or effect of materially distorting the behavior of that person or of a person belonging to that group in a manner that causes or is likely to cause that person or another person significant harm;	2. Any AI system that takes advantage of any of the vulnerabilities of a person or a certain group of persons due to their age or physical or mental disability to substantially alter the behavior of a person belonging to that group in a manner that causes or is likely to cause physical or psychological harm to that person or to another.
Biometric categorization to classify people according to sensitive or protected characteristics	No correspondence.	No correspondence.	b) the placing on the market, putting into service or, use of biometric systems that categorize individually natural persons based on their attributes or sensitive personality characteristics or that which are protected by or based on said attributes.	No correspondence.

			This prohibition does not apply to AI systems intended to be used for therapeutic purposes approved on the basis of the specific informed consent of the persons exposed to them or, where appropriate, their legal guardian.	
Social Score	III – by the public authorities, to evaluate, classify or rank natural persons, based on their social behavior or personality attributes, by means of universal scoring, for access to goods and services and public policies, in an illegitimate or disproportionate manner.	c) The placing on the market, putting into service or use of AI systems for the evaluation or classification of natural persons over a certain period of time based on their social behavior or known or inferred personality characteristics, with the social score leading to either or both of the following: (i) detrimental or unfavorable treatment of certain natural persons or whole groups thereof in social contexts that are unrelated to the contexts in which the data was originally generated or collected; (ii) detrimental or unfavorable treatment of certain natural persons or groups thereof that is unjustified or disproportionate to their social behavior or its gravity;	c) the placing on the market, putting into service or use of AI systems for the evaluation or classification of natural persons or groups thereof over a certain period of time based on their social behavior or known, inferred, or predicted personal or personality characteristics, with the social score leading to either or both of the following consequences: (i) detrimental or unfavorable treatment of certain natural persons or whole groups thereof in social contexts that are unrelated to the contexts in which the data was originally generated or collected; (ii) detrimental or unfavorable treatment of certain natural persons or groups thereof that is unjustified or disproportionate to their social behavior or its gravity;	3. That used by or on behalf of public authorities to assess or classify the trustworthiness of natural persons over a given period of time based on their social behavior or known or predicted personal or personality characteristics, so that the resulting social classification in one or more of the following situations: a. Detrimental or unfavorable treatment of certain persons or whole groups thereof in social contexts that are unrelated to the contexts in which the data was originally generated or collected; b. Detrimental or unfavorable treatment of certain persons or groups, unjustifiably or disproportionately to their social to their social behavior or its gravity.

Identificação biométrica à distância em tempo real (contínua)	Article 15. Within the scope of public security activities, the use of remote biometric identification systems is only allowed, on a continuous basis in spaces accessible to the public, when there is a provision in a specific federal law and judicial authorization in connection with the activity of individualized criminal prosecution, in the following cases: I – prosecution of crimes punishable by a maximum term of imprisonment of more than two years; II – searching for victims of crimes or missing persons; or III – in cases of *being caught red-handed* (caught in the act).	d) The use of 'real-time' remote biometric identification systems in publicly accessible spaces by or on behalf of law enforcement authorities for law enforcement purposes, unless and to the extent that such use is strictly necessary for one of the following purposes: (i) the targeted search for specific victims (i) the prevention of a specific, substantial, and imminent threat to critical infrastructure, the life, health or physical safety of persons or the prevention of terrorist attacks; (iii) the localization or identification of a person suspected of having committed a criminal offence, or the purposes of conducting a criminal investigation, prosecution or executing a criminal penalty for offences (...) and punishable in the Member State concerned by a custodial sentence or detention order for a maximum period of at least three years, or other specific offenses punishable in the Member State concerned by a custodial sentence or detention order for a maximum period of at least five years, as determined by the law of that Member State.	d) The use of 'real-time' remote biometric identification systems in publicly accessible spaces	4. The use of 'real-time' remote biometric identification systems in publicly accessible spaces, unless and to the extent that such use is strictly necessary to achieve one or more of the following objectives: a. The selective search for possible specific victims of a crime, including missing minors. b. The prevention of a specific, significant, and imminent threat to the life or physical safety of persons or of a terrorist attack. c. The detection, location, identification, or prosecution of a person who has committed, or is suspected of having committed, any of the crimes provided for in the Penal Code. The exemptions considered in paragraph 4 of this article shall be subject to an order issued by a Court of Justice and may only be applied by the *Carabineros* of Chile and the Investigative Police.

Assessment of crime risk or recidivism	No correspondence.	No correspondence.	d-A) the placing on the market, putting into service for this specific purpose, or use of an AI system for making risk assessments of natural person or groups of persons in order to assess their risk of committing an offence or recidivism or potential crime or administrative offence based solely on the profiling r on assessing their personality traits and characteristics; including the location of the person, or past criminal behavior of people or groups of people;	No correspondence.
Creation or expansion of databases	No correspondence.	No correspondence.	dB) the placing on the market, putting into service for this specific purpose, or use of an AI system that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage;	No correspondence.
Inference of emotions for certain contexts	No correspondence.	No correspondence.	d-C) the placing on the market, putting into service for this specific purpose, or use of AI systems to infer emotions of a person in law enforcement, border management, the workplace, and educational institutions.	No correspondence.

Subsequent remote biometric identification systems for analyzing recorded images of public spaces	No correspondence.	No correspondence.	dd) The putting into service or use of AI systems for the analysis of recorded images of publicly accessible spaces through subsequent remote biometric identification systems, provided that they are subject to a pre-judicial authorization under EU law and are strictly necessary for the targeted search related to a specific serious criminal offence, as defined in Article 83(1) TFEU, which has already been carried out for law enforcement purposes.	No correspondence.

In light of the above, there is convergence on the ex-ante prohibition of certain uses of AI that present such serious potential risks that they trigger a precautionary approach and, consequently, stricter regulatory intervention. However, there are nuances regarding some of the unacceptable uses, as in the case of biometric identification systems, especially facial recognition, used for the purposes of criminal prosecution. In this case, there is still no consensus on whether there should be a total prior ban, as proposed by the European Parliament, or whether there should be a moratorium allowing their use only in exceptional cases, defined by law.

a.3.2) *High/Elevated Risk*

The high/elevated risk level triggers a more significant level of regulatory intervention, but it is not prohibitive since it authorizes the use of the technology subject to compliance with certain obligations. As a rule, the risk classification technique occurs in a bipartite manner, through the creation of an exemplary list with labeling of examples and/or the establishment of quantitative and qualitative criteria for other activities to be classified as such.

Regarding Bill 2338/23, as well as the proposed European regulation, the Canadian Artificial Intelligence and Data Act (AIDA) and the Chilean Bill 15869/19, the high risk level indicates artificial intelligence systems in which there will be significant regulatory intervention, but which are not prohibited, as was previously illustrated by the pyramid of risks.

However, none of these regulations define what a high-risk artificial intelligence system is, although the proposals provide examples of such systems. As regards the AI Legal Framework stipulated by Bill 2338/23, the categorization of risks related to AI systems is based on a preliminary assessment, as defined in article 13. Article 17 of this same bill lists some AI systems considered to be high risk, based on their purposes, as shown in the table below. This list is exemplary and not exhaustive of the hypotheses of high-risk systems. In order to classify a system as high-risk, the regulations use two methods: (i) if one of the system's purposes is listed in Article 17; and (ii) through analysis based on the quantitative and qualitative criteria set out in Article 18, which deals with the updating of the list of high-risk AI systems by the competent authority.

The hypotheses of high-risk systems in the proposal for a European regulation (European Parliament version of June 2023) are set out in Article 6, complemented by Annex III, identifying two main categories of high-risk AI systems: (i) AI systems intended for use as safety components of products subject to ex ante conformity assessment by third parties; (ii) other autonomous AI systems with implications mainly for fundamental rights, explicitly listed by their area of activity in Annex III. Unlike the European Commission's original proposal, these options are considered high risk by the European Parliament's version, provided that an additional requirement is met: the existence of a significant risk to the health, safety, or fundamental rights of individuals, which would be defined by the European Commission at least six months before the entry into force of the regulation, after public consultation with the AI Authority and other interested parties.

As with Bill 2338/23, the European proposal also provides for the possibility of updating the list of high-risk AI use cases, since it presents a limited number of AI sys-

tems whose risks have already come to fruition or are likely to do so in the near future. In this respect, to ensure that the regulation is constantly updated, according to Article 7, the European Commission can increase, modify, or remove hypotheses, based on certain qualitative and quantitative criteria. As for the Canadian proposal (AIDA), which is still in the process of being drafted, the criteria for designating high-risk systems will be defined in the regulation, which must be aligned with the idea of interoperability with other evolving international regulations on AI, such as the EU AI Act, the AI Principles of the Organization for Economic Cooperation and Development (OECD) and the Risk Management Framework (RMF) of the US National Institute of Standards and Technology (NIST). The supplementary document to the proposal outlines the main analysis factors for determining whether an AI system is high-risk, as well as stressing the importance of paying attention to the capabilities and contexts in which AI systems are used in order to designate the degree of risk, and therefore lists examples of systems that are of interest to the Canadian government in terms of their potential high impacts.

Also in the Canadian context, the Directive on Automated Decision-Making, despite not expressly providing for the category of high risk, establishes impact level III that can be associated with this category, since it is related to automated decisions that will often lead to impacts that may be difficult to reverse and are continuous. In this context, the directive, in its Appendix B, provides that this level is linked to a decision that is likely to have high impacts on the rights of individuals or communities; the equality, dignity, privacy and autonomy of individuals; the health or well-being of individuals or communities; the economic interest of individuals, entities, or communities; and the ongoing sustainability of an ecosystem.

Chile's Bill 15869/19 also includes a list of examples of high-risk AI systems, similar to the EU AI Act, as it associates the level of risk with the intended context of application and provides for the inclusion of more hypotheses to this list in cases involving risks that could harm health, safety or have negative repercussions on fundamental rights. The Chilean initiative differs, however, in that all developers, suppliers and users of AI systems are required to request authorization from the National AI Commission before starting to develop, market, distribute and use these systems in Chilean territory, which means that the level of risk and compliance with the obligations are assessed by the Commission beforehand and even after approval, if the system undergoes substantial modifications. In the other proposals, risk assessment and supervision by the authorities takes place at a later stage.

Regulation	List of high-risk activity examples
Bill 2338/23	AI systems for the following purposes: I - application as security devices in the management and operation of critical infrastructures; II - education and professional training; III - recruitment, screening, filtering, assessment of candidates, decision-making on promotions or termination of employment relationships; IV - evaluation of criteria for access, eligibility, granting, review, reduction, or revocation of private and public services deemed essential; V - assessment of the creditworthiness of natural persons or establishment of their credit rating; VI - dispatching or setting priorities for emergency response services, including firefighters and medical assistance; VII - administration of justice, including systems that assist judicial authorities in investigating facts and enforcing the law; VIII - autonomous vehicles; IX - applications in the healthcare sector, including those aimed at assisting in diagnoses and medical procedures; X - biometric identification systems; XI - criminal investigation and public security; XII - analytical study of crimes relating to natural persons; XIII - investigation by administrative authorities to assess the credibility of evidence during the investigation or repression of offenses, to predict the occurrence or recurrence of an actual or potential offense based on the profiling of individuals; or XIV - migration management and border control.
EU AI Act (European Commission)	AI systems intended for use as a safety component of a product or other AI, other than systems used in one of the following areas (Annex III): I – biometric identification and categorization of natural persons; II – management and operation of critical infrastructures; III – education and professional training; IV – employment, worker management and access to self-employment; V – access to and enjoyment of private services and essential public services and benefits; VI – maintaining public order; VII – managing migration, asylum, and border control; VIII – administration of justice and democratic processes.
EU AI Act (European Parliament)	AI system intended for use as a safety component of a product or other AI, other than systems used in one of the areas below (Annex III), provided that they meet the requirement of posing a significant risk of harm to the health, safety, or fundamental rights of persons or to the environment:

EU AI Act (European Parliament)	I – Biometrics and biometric-based systems: (a) systems used for biometric identification, with the exception of cases prohibited by Article 5; (b) systems used to make inferences about personal characteristics based on biometric data, including emotion recognition (with the exception of prohibited cases); II – Management and operation of critical infrastructures; III – Education and professional training; IV – Employment, management of workers, and access to self-employment; V – Access to private services and essential public services and benefits, such as healthcare, housing, electricity, heating/cooling, and internet; VI – Maintenance of public order; VII – Management of migration, asylum, and border control; VIII – Administration of justice and democratic processes.
AIDA Canadian	The list of examples is currently under development, but already includes: - Screening systems that affect access to services or employment; - Biometric systems used for identification and inference; - Systems that can influence human behavior on a large scale; - Systems critical to health and safety.
Chilean Bill	AI systems used in the following areas: 1. Remote biometric identification in real time or at a later date of people in private spaces. 2. Use in the management of water, electricity, and gas supplies. 3. The allocation and determination of access to educational establishments and the assessment of pupils. 4. Selecting and hiring people for jobs. 5. The assignment of tasks and the monitoring and evaluation of workers' performance and behavior. 6. The assessment of people for access to public assistance benefits and services. 7. Assessing people's creditworthiness or establishing their credit rating. 8. Application in emergency and disaster situations, such as sending or setting priorities for dispatching intervention services (e.g., fire departments or ambulances). 9. Its use to determine the risk of individuals committing crimes or repeating their commission, as well as the risk to potential victims of crimes. 10. Its use at any stage of the investigation and interpretation of facts that may constitute a crime within the scope of a trial. 11. Its use for migration management, asylum, and border control. 12. Likewise, high-risk AI systems will be classified as those that present the risk of causing harm to health and safety, or the risk of having negative repercussions on fundamental rights, whose severity and likelihood are equivalent to or greater than the risks. Negative repercussions associated with the AI systems indicated in the first paragraph of this article.

Regulation	List of qualitative and quantitative criteria for defining high-risk activities
Bill 2338/23	Criteria for assessing whether a system is high risk: I – implementation on a large scale, taking into account the number of people affected and the geographical extent, as well as its duration and frequency; II – the system may negatively impact the exercise of rights and freedoms or the use of a service; III – the system has a high potential for material or moral damage, or is discriminatory; IV – the system affects people from a specific vulnerable group; V – the possible harmful results of the artificial intelligence system are irreversible or difficult to reverse; VI – a similar artificial intelligence system has previously caused material or moral damage; VII – a low degree of transparency, explainability and auditability of the artificial intelligence system, which makes it difficult to control or supervise; VIII – high level of identifiability of data subjects, including the processing of genetic and biometric data for the purpose of uniquely identifying a natural person, especially when the processing includes combining, matching, or comparing data from several sources; IX – when there are reasonable expectations by the affected party regarding the use of their personal data in the artificial intelligence system, in particular the expectation of confidentiality, such as in the processing of sensitive or confidential data (Article 18)
EU AI Act (European Commission)	To be considered high-risk, the system must meet two requirements: (a) Be used in one of the 8 areas described in Annex III; and (b) It presents a risk of harm to health, safety, or the creation of adverse effects on fundamental rights: in terms of severity and probability of occurrence, it presents a greater or equivalent risk to the risks of harm, or adverse effects created by the high-risk systems already listed in Annex III. In order to conduct this analysis, the following criteria are analyzed: - The intended purpose of the AI system; - The extent to which the system will be used or is expected to be used; - The extent to which the use of an AI system has already caused harm to health and safety or adverse impact on fundamental rights or given rise to significant concerns regarding the materialization of such harm or adverse impact, as demonstrated by documented reports or allegations submitted to competent national authorities; - The potential extent of such damage or adverse impacts, especially in terms of their intensity and their capacity to affect a plurality of people; - The extent to which people who are potentially harmed or negatively affected depend on the outcome produced, since, for practical or legal reasons, it is not reasonably possible to opt out of this outcome; - The extent to which potentially harmed or negatively affected people are in a vulnerable position in relation to the user of an AI system, for example, due to an imbalance of power, knowledge, economic or social circumstances or age;

	- To what extent the result produced with an AI system is easily reversible, so results that have an impact on people's health or safety should not be considered easily reversible; - To what extent does current EU legislation provide for: (i) effective remedies in relation to the risks posed by an AI system, excluding claims for compensation; (ii) effective measures to prevent or substantially minimize those risks. (Article 7)
EU AI Act (European Parliament)	For a system to be considered high-risk, it must pose a significant risk of harm to health and safety, or an adverse impact on fundamental rights, the environment, or democracy and the rule of law. This risk is, in terms of its severity and likelihood of occurrence, equivalent to or greater than the risk of harm or adverse impact posed by Annex III high-risk AI systems. In order to conduct this analysis, the following criteria are analyzed: - The AI system's intended purpose; - The general capabilities and functionality of the system, regardless of its intended purpose; - The extent to which the system will be used or is expected to be used; - The nature and quantity of data processed and used by the system; - The extent to which the AI system acts autonomously; - The extent to which the use of an AI system has already caused harm to health and safety, had an adverse impact on fundamental rights, the environment, democracy, and the rule of law, or given rise to significant concerns regarding the likelihood of such harm or adverse impacts; - The potential extent of such harm or adverse impacts, especially in terms of their intensity and their capacity to affect a plurality of people or disproportionately affect a specific group of people; - The extent to which persons potentially harmed or adversely affected are dependent on the result produced, and that result is purely incidental with respect to that which is relevant, in particular because, for practical or legal reasons, it is not reasonably possible to opt out of that result; - The potential misuse and malicious use of the AI system and the technology that supports it; - The extent to which there is an imbalance of power or potentially harmed or negatively affected people are in a vulnerable position in relation to the user of an AI system, for example due to status, authority, knowledge or economic, social or age circumstances; - The extent of the availability and use of effective technical solutions and mechanisms for the control, reliability, and rectification of the AI system; - The magnitude and likelihood of benefits from the implementation of the AI system for individuals, groups, or society in general, including possible improvements in product safety; - The extent of human supervision and the possibility of a human being interfering to override a decision or recommendations that could cause potential harm;

	- To what extent the result produced with an AI system is easily reversible, so results that have an impact on people's health or safety should not be considered easily reversible; - To what extent does existing EU legislation provide for: (i) effective remedies for damage caused by an AI-system, excluding claims for direct or indirect damage; (ii) effective measures to prevent or substantially minimize such risks. (Art. 7º)
AIDA **Canadian**	The Canadian government lists the following criteria to determine which AI systems would be considered high impact: - Evidence of health and safety risks, or risk of adverse impact on human rights, based on both the intended purpose and possible unintended consequences; - The severity of the potential damage; - The scale of use; - The nature of the damage or adverse impacts that have already occurred; - The extent to which, for practical or legal reasons, it is not reasonably possible to opt out of that system; - Imbalances in economic or social circumstances, or age of those affected; and - The degree to which the risks are adequately regulated by other law
Canada: Directive on Automated Decision-Making + Algorithmic Impact Assessment tool	The Directive addresses AI in the form of automated decisions. In this context, high-impact automated decisions are those that affect: individual or community rights; equality, dignity, privacy, and autonomy; the health and well-being of individuals and groups; and the ongoing sustainability of an ecosystem. The criteria for identifying whether decisions have a high impact are: (i) the difficulty of reversing their results; (ii) whether their results are continuous.
Chilean Bill	Criteria for assessing whether an AI system is high risk: - If it presents risks of causing harm to health and safety or the risk of having negative repercussions on fundamental rights, the severity and likelihood of which are equivalent to or greater than the risks of harm or negative repercussions associated with the list of examples of high-risk AI systems in Article 4. (Article 4)

There is a certain convergence in AI bills around the world in proposing a list of examples of high-risk AI systems, the possibility of updating them and the inclusion of new systems (not originally listed) based on the definition of qualitative and quantitative criteria. These criteria are better developed in certain projects, such as Bill 2338/2023 and in the versions of the EU AI Act, but it is important that they are defined so that the legislation regulating AI is not static and can survive the passage of time and the rapid advances in technology.

Despite their convergence, there are still topics up for discussion, such as the clas-

sification of AI systems used for biometric identification, and the use of real-time remote biometric identification systems in publicly accessible spaces, categorized as a prohibited practice by the European Parliament's version of the EU AI Act. In the other proposals, such as Bill 2338/23, the EU Council's version of the European proposal and the Chilean Bill, the ban only applies to the use of these systems for public security purposes, with some exceptions listed.

a.3.3) *Low risk (Residual)*

The categorization of low or moderate risk represents the least burdensome level of regulatory intervention, which means that the related obligations are more lenient for AI agents. This classification can be achieved using qualitative and quantitative identification elements or by defining it as a residual category, which can be followed by examples, either in the text of the law itself or in the explanatory memorandum. In the latter case, there is no provision for direct criteria, but indirect classification through a legislative technique of exclusion - in other words, cases that do not fall within the other more strict levels of risk are qualified as residual.

In some regulations, such as Bill 2338/2023, low risk represents a residual category, as it is made up of all AI systems that fall outside the classification of excessive/unacceptable risk or high risk and is not expressly mentioned in the text of the law. In these cases, the bill provides for governance structures and internal processes to be mandatory for all AI systems, including those at the residual level, in order to guarantee the safety of the systems and the fulfillment of the rights of those affected. These obligations are listed in Article 19 and include, for example: (i) transparency measures regarding the use of artificial intelligence systems in interaction with natural persons, which includes the use of appropriate human-machine interfaces that are sufficiently clear and informative; (ii) transparency regarding the governance measures adopted in the development and use of the artificial intelligence system by the organization; (iii) appropriate data management measures for the mitigation and prevention of potential discriminatory biases; among others.

For the European Union's proposal, for example, the explanatory memorandum states that the heaviest regulatory burdens will only be imposed when an AI system is likely to present high risks to fundamental rights and safety. For other non-high risk AI systems, considered to be a limited risk category, the proposal only imposes transparency obligations, such as the provision of information to signal the use of an AI system that interacts with human beings. However, it is worth noting that Article 52 (4) states that these transparency obligations also apply to other high-risk systems, which fall under

Title III of the proposal.

Furthermore, although not expressly mentioned in the text of the proposed regulation, some studies led by the European Commission and the European Parliament also mention a low or minimum risk level, which would not be subject to extra obligations under the proposal and could be developed and used in the EU without fulfilling any additional legal obligations. However, the proposal provides for the creation of codes of conduct to encourage providers of non-high risk AI systems to voluntarily apply the mandatory requirements for high risk AI systems[63].

Regulation	Residual risk or list of examples?	Examples	Obligations
Bill 2338/23	Residual risk (assumed - no explicit mention)	-	Governance structures defined in Article 19, including, for example: transparency measures, data management to mitigate and prevent discriminatory bias and information security from the design to the operation of the system.
PL 2338/23	Limited risk - examples	Systems that interact with humans (i.e., chatbots), emotion recognition systems, biometric categorization systems and AI systems that generate or manipulate image, audio, or video content (i.e., deepfakes)	Only minimal transparency obligations
	Residual (low or minimal)	-	No obligations.
AIDA	Assumed residual risk (not explicitly mentioned)	-	No obligations.

As this level is less risky for fundamental rights, it receives less attention in the documents analyzed. As a rule, most of the regulatory proposals include low risk as a residual category, based on the classification by exclusion of all AI systems not classified in the more intense risk levels. However, even though this is a category which is less

63 https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf.

focused on by regulation, there are still provisions for minimum obligations, especially those related to transparency.

a.4) *Ex-ante* and *ex-post* risk approach

As the risk-based approach is expanded when it comes to AI regulation, the activity of risk classification also becomes necessary, as seen in the previous topics. As mentioned by Hood et al (2001), the regulation of risk - and consequently its classification - varies from one domain to another and can change over time[64]. The OECD principles on AI require that agents are responsible for the proper functioning of their AI systems, according to their role, context, and capacity for action[65]. Added to the fact that AI is a complex and rapidly evolving technology, such a risk classification exercise in this scenario is not trivial and always requires a contextual analysis.

Within the AI field, in addition to being extremely contextual, risk analysis can take place beforehand (*ex-ante*) or after the fact (*ex-post*). In the first case, the assessment can take place either by creating a list of examples of unacceptable and high risks (or by creating criteria for their classification), but also by conducting an assessment of AI systems from the product and/or service design phase, which must be allocated to the chain of agents involved. For Brazil's Bill 2338/23, such an analysis is provided for in Article 13, which stipulates that every AI system must undergo a preliminary assessment to classify its degree of risk - which is fundamental for allocating obligations to each of the actors.

According to Kaminski, risk regulation usually focuses on ex-ante measures and underutilizes post-marketing tools[66]. However, the risk-based approach with its ex-ante and ex-post evaluation makes the risk-based regulatory process more complete, by means of a learning movement, since it allows for the reclassification of systems at a later date, should there be any significant changes during the course of its application. This means that regulation takes place in a dynamic, non-static and collaborative way. This approach is found, for example, in the European Union (EU AI Act in all its versions), in Brazil (Bill 2338/23), in Chile (Bill 15869-19) and in Canada with the provision of criteria for reclassification of the list of examples of unacceptable and high risks, as well as the creation of a publicly accessible database of AIs, which is generally mandatory for high-risk cases - which allows the participation of the whole society in monitoring and evaluating these risks after the systems have been implemented.

64 HOOD *et al*, 2001, p. 3.

65 OECD. OECD AI Principles overview. Disponível em: https://oecd.ai/en/ai-principles.

66 KAMINSKI, 2022, p. 72–73.

Regulation	*Ex-ante* previsions	*Ex-post* previsions	Expected creation of a publicly accessible database of high-risk AIs
Bill 2338/23	Yes	Yes	Yes
European Proposal (*AI Act*)	Yes	Yes	Yes
Chilean Bill 15869-19	Yes	Yes	Yes
Canada's Algorithmic Impact Assessment Tool	Yes	Yes	Yes

Therefore, as much as risk regulation varies according to the area in which it is applied, it is certain that its analysis and risk classification must always involve contextual action, which can be done either ex-ante or ex-post the implementation of AI systems. Prior and subsequent risk regulation and classification is generally expressed in impact assessment tools, which, as will be addressed in the next topic, must be continuous, updatable from time to time (or in the event of significant changes to the systems) and with significant public participation from all sectors of society.

AXES 2 – Algorithmic Impact Assessment (AIA)

Impact assessments, which are well-known in the environmental and personal data protection fields in terms of environmental impact reports and personal data protection impact reports, are governance tools that have emerged to analyze the possible consequences of an initiative on relevant social interests. Based on this analysis, they support an informed decision-making process on whether the initiative should be conducted and, if so, under what conditions. They are applied in situations where there is uncertainty about future events, such as the emergence of new technologies[67]. For this reason, impact assessments are mechanisms for generating evidence for decision-making and for protecting certain societal concerns[68].

From the outset, it should be noted that the impact assessment tool differs from other organizational activities, such as a regulatory compliance assessment (which can be done ex-post, although this is not ideal), due to its precautionary and preventive na-

67 KLOZA *et al*, 2019.

68 KLOZA *et al*, 2017.

ture. Its aim is to identify risks and apply efficient mitigation measures before implementing a given technology, following public scrutiny that unleashes social control and a governance network[69].

For this reason, one of the main objectives at stake is to ensure that impact assessments - from the environmental field, as in the case of assessing the construction of a highway next to the riparian forest of a river, to the field of data protection - create a procedure in which all interested parties can understand and influence a given decision-making process. It's a question of procedural justice[70], and what's at stake is not only a fair outcome, but also whether the process used to achieve that outcome is fair.

Given the growing application of artificial intelligence systems to automate decisions in our daily lives, this is related to what has come to be called "due process of information"[71]. In other words, ensuring the right to a fair hearing and a full defense and, consequently, curbing actions that unduly interfere with public freedoms - e.g., predictive policing - and individual rights - e.g., freedom of expression in the content moderation scenario - through greater control over the procedures that are conducted.

Furthermore, in conducting the impact assessment tool, it should not be seen as a burden or mere obligation for the supplier, but as an opportunity. Given the nature of AI products/services and their resources and scale, the proposed evaluation model can significantly help companies and other entities to develop human-centered and effective AIs, even in challenging contexts[72]. As a result, confidence is generated not only in technology, but also in the economic exchanges around it.

Finally, it is essential to stress that, more important than merely providing for an algorithmic impact assessment tool, is that it be minimally proceduralized so that it becomes an effective tool for due process and accountability. By way of example, the Brazilian Bill 21/20 conceptualized what an artificial intelligence impact report would be in Article 2, item VI, but failed to provide greater detail as to its objectives, deadlines, and minimum parameters, which leads to legal uncertainty. Therefore, in addition to the legal provision for conducting an AIA, it is essential that there is also a definition of minimum parameters for methodology, criteria, stages and, possibly, provisions on the need for publication and periodic review. In this respect, Bill 2338/2023 is a step forward compared to the others, similar to what is done in the AI EU Act, Canada, and other international instruments.

69 BIONI, Bruno Ricardo. *Regulation and Protection of Personal Data – The Principle of Accountability*. São Paulo: Editora Forense, 2022. 320p.

70 KLOZA *et al*, 2019.

71 BIONI; MARTINS, 2020.

72 MANTELERO, 2022.

b.1) Methodology, criteria, and timing of implementation

As a preliminary lesson drawn from the Personal Data Protection Impact Report[73], the lack of minimum proceduralization, i.e., systematization of what this tool should contain (such as deadlines, criteria and chosen methodology), hinders its implementation. At the same time, this type of parameterization should not be too prescriptive so as not to stiffen a tool that should be as dynamic as the development of AI technologies and other data processing techniques. Consequently, future AI regulations are welcome to present a minimal systematization, a foundation for a sound building of algorithmic impact assessment. In other words, a floor and not a ceiling for modeling this important tool.

A preliminary study on AIA conducted by the Council of Europe's Ad Hoc Committee on Artificial Intelligence found that[74], when assessing the general frameworks of human rights impact assessments, there is a tendency for these instruments to focus on the adverse impacts of a given initiative on these rights, which is also the case with most of the current impact assessment models for AI systems.

According to CAHAI, AI impact assessments should be developed according to this approach, although the use of AI does not only generate adverse impacts, as the technology has many advantages and can create a huge beneficial impact for humanity. Nevertheless, according to CAHAI[75], the primary specific function of human rights impact assessments, which should underpin AIA, should be to detect possible risks of human rights violations arising from a given AI system, and not to balance them against possible beneficial impacts arising from such an application.

This means that the trade-off between benefits and risks would not necessarily be part of the impact assessment methodology but would subsequently help in assessing whether or not to implement AI systems. For example, in certain cases, decision-makers,

73 BIONI, Bruno Ricardo; ZANATTA, Rafael A. F.; RIELLI, Mariana Marques. Contribution to the Public Consultation on the Brazilian Artificial Intelligence Strategy, Data Privacy Brazil Research, available at: https://www.dataprivacybr.org/wp-content/uploads/2020/06/E-BOOK-CONTRIBUIC%CC%A7A%CC%83O-DPBR-INTELIGE%CC%82NCIA-ARTIFICIAL-FINAL.pdf; BIONI, Bruno; EILBERG, Daniela Dora; CUNHA, Brenda; SALIBA, Pedro; VERGILI, Gabriela. Data protection in the criminal and public security field: technical note on the Draft Data Protection Law for public security and criminal investigation. São Paulo: Data Privacy Brazil Research Association, 2020.

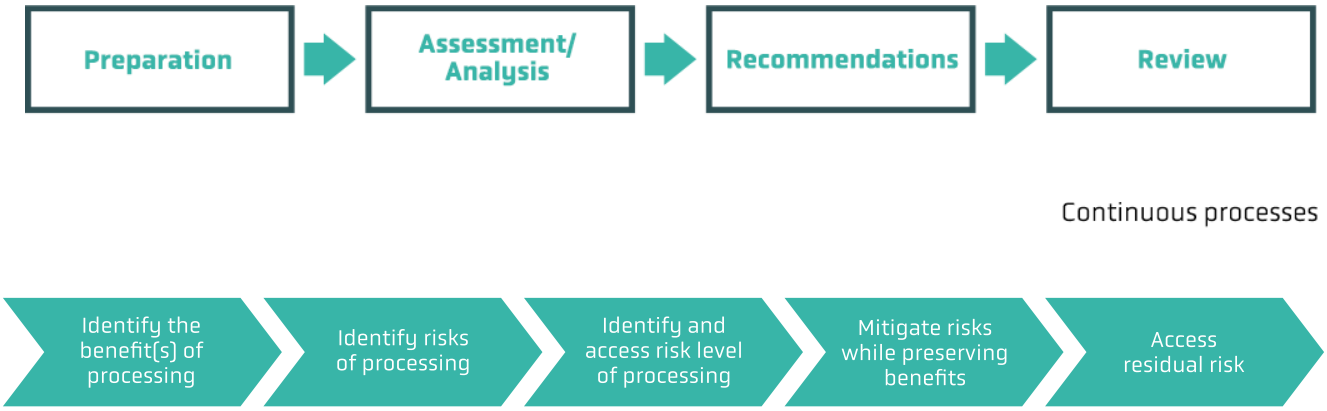
74 CAHAI, 2022, p. 4.

75 *"Obviously, this does not imply that the use of AI generates adverse impacts only. AI has many advantages and can create a huge beneficial impact for mankind. It may even assist in the enjoyment, protection and strengthening of human rights, and this positive contribution should not be neglected. However, the specific function of HRIA is to detect possible risks of infringement for human rights arising from a given AI system, and not to balance them against possible beneficial impacts arising from such an application. Balancing benefits against risks is not part of the assessment methodology but would rather be performed later as part of a judgement of opportunity as to whether deploy such application. For instance, in certain cases public authorities could conclude that the beneficial impacts offset adverse impact and hence decide using such application for a given purpose. If in this case one or more human rights are curbed (which the HRDRIA can help assess) it is essential that this occurs in a manner that is justified through an approach that is both proportionate and necessary in a democratic society, for instance in the interest of national security or another legitimate public interest".* CAHAI. Human Rights, Democracy and Rule of Law Impact Assessment of AI systems. Policy Development Group (CAHAI-PDG). Strasburg: May 21, 2021. p. 4.

such as public authorities, may conclude that the beneficial impacts outweigh the adverse impacts and therefore decide to use such an application for a particular purpose.

Under the UNESCO Recommendations on the ethics of AI systems, the role of the AIA tool would be to "identify and assess the benefits, concerns and risks of AI systems, as well as risk prevention, mitigation and remediation and monitoring measures," based on the impacts on human rights and fundamental freedoms, especially the rights of people in precarious and vulnerable situations, labor rights and the environment[76].

In addition, from a comparative analysis of impact evaluations in different areas conducted by d.pia.lab, it is possible to establish elements that constitute good practice for impact evaluations that can be adapted to different areas[77]. To this end, a generic method for impact assessment has been established, consisting of 10 steps grouped into 5 phases. These include: (i) preparation; (ii) evaluation/analysis; (iii) recommendations; (iv) ongoing steps; (v) review[78]. When transferred to the context of AI, it is understood that the tool serves to identify, describe, and analyze both the possible consequences of the system under analysis and possible solutions to address those consequences.



Similarly, CAHAIs "Human Rights, Democracy and Rule of Law Impact Assessment of AI Systems" report defined four minimum steps for the development of this tool, so as to include the identification of relevant rights, assessing the impacts on these rights (including, as criteria, the scope and scale of the application and the potential number of

76 UNESCO. Recommendation on the Ethics of Artificial Intelligence. Implemented on November 23, 2021, and published in 2022. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000381137.

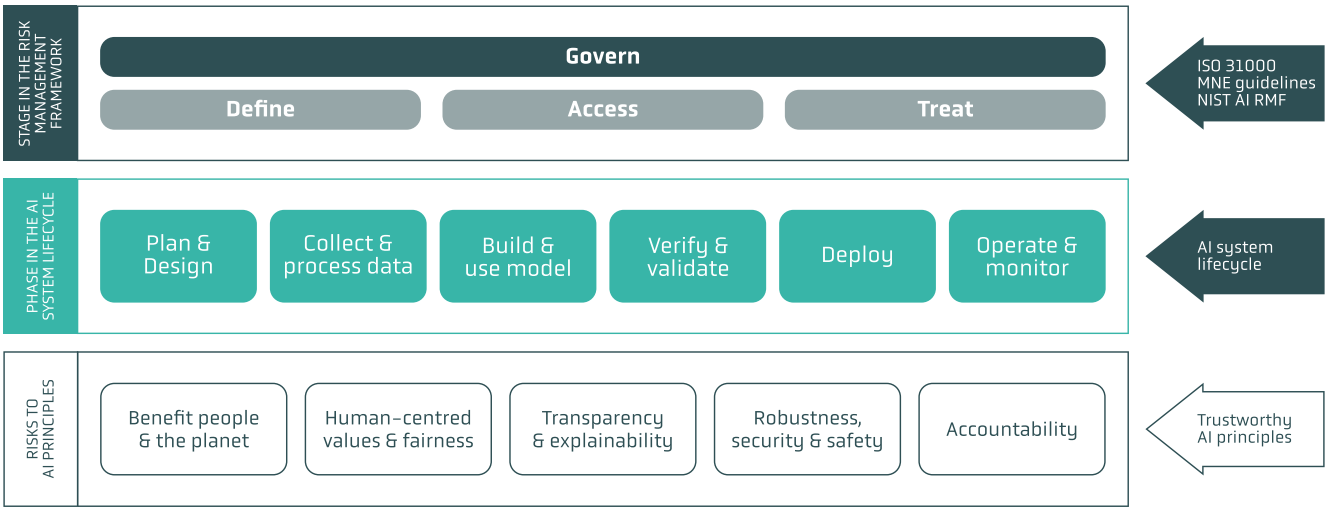
77 KLOZA, D., *et al.* (2017). Data protection impact assessments in the European Union: complementing the new legal framework towards a more robust protection of individuals. **d.pia.lab Policy Brief**, (1/2017), 1–4. https://doi.org/10.31228/osf.io/b68em, https://doi.org/10.5281/zenodo.5121575.

78 KLOZA, D., *et al.* (2019). Towards a method for data protection impact assessment: Making sense of GDPR requirements. d.pia.lab Policy Brief, 1(2019), 1–8. https://doi.org/10.31228/osf.io/es8bm, https://doi.org/10.5281/zenodo.5121534.

people impacted), governance mechanisms and ongoing evaluations[79].

A report on AI accountability produced by the OECD[80] n 2023 also highlighted the existence of at least 4 stages for risk management of AI systems, which would consist of: definition (scope, context, actors involved and analysis criteria), assessment (identification of individual and collective risks based on severity x probability), risk treatment (mitigation measures) and governance (monitoring and review). This guidance is in line with other international frameworks from, for example, NIST, ISO 31000 and other OECD documents[81].

Recently, the OECD also published "Common guideposts to promote interoperability in AI risk management," in which it confirms that the main risk management models and frameworks align with these four steps. Although the target audience, the risk scope, the segment of the AI lifecycle, the specific terminology used and the order of the steps themselves may vary between existing documents, the models generally seek to achieve the same results (e.g., responsible, ethical, and trustworthy AI) through a similar four-step risk management processes[82].



(OECD, 2023b)

79 CAHAI. Human Rights, Democracy and Rule of Law Impact Assessment of AI systems. Conselho da Europa, CAHAI-PDG (2021)5. Strasbourg, May 21, 2021. Available at: https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3.

80 OECD. Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI. Published on Feb. 23, 2023. Available at: https://www.oecd-ilibrary.org/science-and-technology/advancing-accountability-in-ai_2448f04b-en.

81 OECD, 2023a.

82 OECD. Common guideposts to promote interoperability in AI risk management. 07 nov. 2023. Available at: https://www.oecd-ilibrary.org/science-and-technology/common-guideposts-to-promote-interoperability-in-ai-risk-management_ba602d18-en.

Similarly, in Brazil, Bill 2338/2023 also defines a methodology of at least four stages represented by preparation, risk awareness, risk mitigation and monitoring (Article 24).

In this respect, according to Alessandro Mantelero (2022), there are at least three essential factors that must be considered in a risk analysis: (i) identification of this risk; (ii) likelihood of the risk materializing; (iii) severity of the risk identified. For identification, it is recommended that rights be broadly included as potentially affected categories, in order to guarantee the comprehensive protection of natural persons and the different groups impacted by the possible risks triggered by the use of the AI system, as well as the environment. Therefore, as mentioned, when proposing a hybrid model based on risks and rights, the risk relates to the potential damage to the fundamental rights and freedoms of natural persons[83], considering limitations and restrictions, regardless of the material damage.

Thus, formal risk analysis relies on a mathematical problem involving elements such as the probability of an event happening, and the severity of the damage potentially caused by that event[84]. In other words, the expected impact of the identified risks is assessed considering both the probability and the severity of the expected negative consequences, using a variable scale, which is usually conducted in four stages (low, medium, high, extremely high/excessive/unacceptable). This gradation, however, varies depending on the risk matrix adopted, which can have different gradations, ranging from the simplest (three levels: low, moderate, or high) to the most complex (with four to five levels of risk, for example)[85].

83 GOMES, Maria Cecília. Data protection impact report: a brief analysis of its definition and role in the LGPD. Revista da AASP, n. 144, 2019. p. 10–11.

84 KAMINSKI, 2022, p. 8.

85 TV Senado. Commission of jurists promotes debates on regulation of artificial intelligence (part 2) – April 29th,

Mantelero (2022) considers probability as a combination of two elements: the likelihood of adverse consequences and exposure (potential number of people at risk), while severity is assessed considering the nature of the potential damage to the exercise of rights and its consequences, which also includes checking the effort needed to overcome the potential damage and reverse the adverse effects. Along these lines, for example, §1 of Article 24 of Bill 2338/2023, which provides for impact assessment for high-risk systems, requires this tool to consider and record both risk identification - which should encompass both known and foreseeable risks and those that can be expected - and the likelihood and severity of adverse consequences.

As the comparative table below illustrates, the proposed assessments focus on risks to the individual, diffuse, collective and individual homogeneous rights[86] of those affected by AI systems, in a broad way, designed to guarantee comprehensive protection not only of human rights but also of ethical, democratic and rule of law values.

Mantelero (2022) argues that AI systems carry with them a complexity that requires impact assessments to be based on a hybrid model for analyzing their ethical and social impact, along with legal dimensions such as human rights. To this end, he advocates the need for a multi-stakeholder, human-centered approach, combining the universality of human rights with the local dimension of social values. Similarly, CAHAI proposes a comprehensive analysis of human rights, democracy, and the rule of law[87] and UNESCO mentions that the assessment should refer not only to the individuals or groups/communities affected, but also to the environment[88].

Thus, considering that the risks vary depending on the AI system used, and that an impact assessment is not a trivial process, both in terms of carrying it out and analyzing it, this requirement is generally restricted to high-risk AI systems, without prejudice

---

2022. Published on April 29, 2022. Speech by Professor Maria Cecília Gomes. Available at: https://www.youtube.com/watch?v=P_yWp-2ZIZs&t=51s. Accessed on July 21, 2023.

86 Collective rights in the broad sense are an important social achievement and were enshrined in the 1988 Brazilian Federal Constitution and in other regulations of the collective (procedural) microsystem, such as the Public Civil Action Law and the Consumer Protection Code. These rights can be divided into diffuse, collective (stricto sensu) and individual homogeneous rights, as provided for in the sole paragraph of Article 81 of the CDC. In this context, the interests or rights are understood to be transindividual, indivisible in nature, owned by indeterminate people linked by factual circumstances, as in the case of the right to a healthy environment. Collective interests or rights in the strict sense, on the other hand, are those that are transindividual, of an indivisible nature, owned by a group, category or class of people linked to each other or to the opposing party by a basic legal relationship, as in the case of consumers of essential public services. In this case, it is possible to determine who the holders are, since there is a legal relationship between the people affected. Finally, homogeneous individual interests or rights are those arising from an event with a common origin, as in the case of consumers injured by a defective product – here, it is possible to file both an individual and a collective action; National Council of Public Prosecutors. Collective Rights Portal. Available at: https://www.cnmp.mp.br/direitoscoletivos/; GAJARDONI, Fernando da Fonseca. **Diffuse and Collective Rights I: General Theory of Collective Proceedings**. São Paulo: Saraiva, 2012.

87 https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3.

88 UNESCO. Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence. Published in 2023. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000386276/PDF/386276eng.pdf.multi.

to it being conducted as good practice for lower-risk AI systems, as will be seen below. This is what is provided for in Bill 2338/23 and the EU AI Act, as well as the Canadian AIDA. However, in addition to the high-risk classification as a criterion for removing the obligation to conduct impact assessments, there are some initiatives that associate other criteria, such as the nature of the organization, for example AI systems used by public authorities, as reinforced by the AI Ethical Impact Assessment Tool provided within the framework of UNESCO's[89] recommendation.

In this regard, considering that the triggering of the obligation to prepare an AIA is associated with the degree of risk of the system in question, it is necessary for there to be a process of evaluation of this degree by the players in the AI production chain. By way of example, Article 13 of Bill 2338/2023 provides that, prior to being placed on the market or used in service, every AI system must undergo a preliminary assessment to classify its risk, according to the criteria defined in the articles relating to excessive and high risk. In this context, the Brazilian project also defines the obligation to record and document this assessment for the purposes of responsibility and accountability a posteriori, preventing the risk classification process from being left solely to the regulated actors, including allowing the competent authority to determine the reclassification of systems and penalties for fraudulent analysis.

As for the European context, the Parliament's version of the AI Act does not expressly provide for a prior assessment but leaves it implicit by providing some criteria for defining unacceptable and high risks, without, however, determining mechanisms for monitoring by the competent authorities, which could end up making it difficult to implement the law efficiently and harmoniously[90].

Furthermore, beyond its scope, it is important to note that the AIA must be considered a continuous process and not an immediate moment in time[91]. A consensus among the documents seems to be that the AIA should be completed before the system is actually made available to the public, either as a service or as a product. These provisions are in line with what Maria Cecília Gomes argued at a public hearing held under the auspices of CJSUBIA in April 2022. According to her, this assessment ought to be conducted at the time the AI is developed and the risks need to be assessed before they become a reality[92]. The AIA tool is therefore inherently preventive, based on a logic of ex-ante

89 UNESCO, 2023.

90 Access Now. EU Trilogues: The AI Act must protect people's rights. Publicado em 12 jul 2023. Disponível em: https://www.accessnow.org/press-release/eu-trilogues-ai-act/; Access Now. Joint statement: EU legislators must close dangerous loophole in AI Act. Published on September 7, 2023. Available at: https://www.accessnow.org/press-release/eu-trilogues-ai-act/https://www.accessnow.org/press-release/joint-statement-eu-legislators-must-close-dangerous-loophole-in-ai-act/.

91 CAHAI, 2021, p. 4.

92 Ibid.

regulation, rather than being a diagnosis of future adverse events[93], when the risk has already materialized after the technology has been made available on the market[94].

The obligation to conduct algorithmic impact assessments, prior to making a service available or placing a product on the market, can bring benefits to organizations based on a shift towards anticipatory and ex ante decision-making. Organizations are now able to reflect on the consequences of their initiatives, as well as the means to minimize or sometimes even avoid negative and unintended consequences before they occur, which leads them to gain public trust in the medium and long term - and consequently, reputational gain[95].

However, in addition to its preventative importance, it is also a tool that consists of a systematic process that begins reasonably early in the life cycle of a single initiative (such as AI systems), before its deployment, continues throughout its life cycle and - as society changes, dangers evolve and knowledge grows - is revisited when necessary[96]. It is therefore possible to refer to impact assessments as a "living instrument"[97] that will continuously influence the development of AI systems.

Accordingly, UNESCO also maintains that the AIA tool should be a dynamic document, completed progressively and iteratively at different stages of the AI lifecycle, including, for example, the following stages: (i) design, development and pre-acquisition of the system; (ii) acquisition, when the AIA tool can help both in the selection of a supplier and in the formulation of contractual obligations; (iii) after the implementation of the system, when the AIA should be reviewed at regular intervals, especially as the responses may change over time as the technology evolves[98].

Bill 2338/2023 expressly provides for periodic updates of the AIA, which must be part of the entire life cycle of high-risk AI systems (Articles 25 and Article 24, fourth paragraph). This means that, even though it is a tool that is predominantly conducted before the technology is launched on the market, it is essential that it is updated throughout the AI lifecycle. Not only because new techniques emerge over time, but also because incidents can inform the updating of the entire risk management process to make it even more resilient. Consequently, there is also an ex-post aspect to this tool.

93 KAMINSKI, 2022, p. 19.

94 https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3.

95 KLOZA, Dariusz *et al* Data protection impact assessments in the European Union: complementing the new legal framework towards a more robust protection of individuals. d.pia.lab Policy Brief, 2017.

96 MANTELERO, 2022.

97 In Kloza et al (2017), the authors refer to impact assessment as a "living instrument" to explain the fact that the tool is constantly in need of reflection, since the assessment of technology starts early (before it is implemented), continues throughout its life cycle (once it has been implemented) and, as society advances, the dangers increase and growth evolves, it needs to be revised in order to influence the design of the technology itself.

98 UNESCO, 2023, p. 8.

Furthermore, in risk-based regulation, it is an instrument inserted into the regulatory learning process, which must be dynamic and iterative. In other words, risk management is not merely an ex-ante verification to be completed, but a hybrid process that must be repeated from time to time and altered as risks and knowledge about technologies change, allowing the risk to be recalibrated ex-post. This is what is observed, for example, in the GDPR, EU AI Act, Bill 2338/23 and other documents listed in the table below.

In this context, there is no single answer as to who should be responsible for triggering the obligation to renew or update the algorithmic impact assessment, nor the ideal frequency at which it should be conducted. However, there seems to be a tendency for it to be reviewed from time to time, especially if the risks or circumstances surrounding the technology change. As an example, Bill 2338/2023 assigns the future competent AI authority to define the frequency of the AIA review, while the bill on automated decision tools from the US state of California defines its annual update.

Legislation	Is there any mention of the impact assessment being an ongoing document?	Excerpt	Is there provision for some kind of state supervision with powers to do so?	Excerpt
Bill 2338/2023	Yes	Article 19, § 1. The governance measures of artificial intelligence systems apply throughout their entire lifecycle, from initial conception to the termination of their activities and discontinuation. Article 24, § 4 The competent authority will be responsible for regulating the frequency for updating impact assessments, taking into account the life cycle of high-risk artificial intelligence systems and fields of application, and may include best sectoral practices. Article 25: The algorithmic impact assessment will consist of a continuous iterative process, conducted throughout the life cycle of high-risk artificial intelligence systems, requiring periodic updates. § 1º The competent authority will be responsible for regulating the frequency with which impact assessments are updated.	Yes	Article 24, § 4º The competent authority will be responsible for regulating the frequency with which impact assessments are updated, taking into account the life cycle of high-risk artificial intelligence systems and fields of application, and may incorporate best sectoral practices. Article 25, § 1º The competent authority will be responsible for regulating the frequency with which impact assessments are updated.
EU AI Act (European Parliament's version)[99]	Yes	Article 9 (1) A risk management system shall be established, implemented, documented, and maintained in relation to high-risk AI	Yes	(85a) Given the rapid technological developments and the required technical expertise in conducting the assessment of high-risk AI

99 The original version of the EU AI Act did not provide for an impact assessment, but only for a risk management process in the case of high-risk AI systems (article 9). However, in the latest version of the proposal, published by the European Parliament, in addition to risk management, a human rights impact assessment must also be drawn up by the deployer for high-risk AI systems (Article 29a).

EU AI Act (European Parliament's version)		systems, throughout the entire lifecycle of the AI system. The risk management system can be integrated into, or a part of, already existing risk management procedures relating to the relevant Union sectoral law insofar as it fulfills the requirements of this article. (2) The risk management system shall consist of a continuous iterative process run throughout the entire lifecycle of a high-risk AI system, requiring regular review, and updating of the risk.		systems, the Commission should regularly review the implementation of this Regulation, in particular the prohibited AI systems, the transparency obligations and the list of high-risk areas and use cases, at least every year, while consulting the AI office and the relevant stakeholders.
GDPR	Yes	Article 35 (11) Where necessary, the controller shall conduct a review to assess if processing is performed in accordance with the data protection impact assessment at least when there is a change of the risk represented by processing operations.	Yes	Article 35 (11) Where necessary, the controller shall conduct a review to assess if processing is performed in accordance with the data protection impact assessment at least when there is a change of the risk represented by processing operations.
Directive on Automated Decision-Making (Canadá)	Yes	6.1.3 Reviewing and updating the Algorithmic Impact Assessment on a scheduled basis, including when the functionality or scope of the automated decision system changes.	Yes	Subject to the necessary delegations, the Chief Information Officer of Canada is responsible for: (...) 8.2 Developing and maintaining the Algorithmic Impact Assessment and any supporting documentation.
Algorithmic Impact Assessment tool (Canadá)	Yes	3.1 When to complete the AIA: The AIA should be completed at the beginning of the design phase of a project. The results of the AIA will guide the mitigation and consultation requirements to be met during the implementation of the automated decision system as per the directive. The AIA should be completed a second time, prior to the production of the system, to validate that the results accurately reflect the system that was built	Yes	3.2 What to consider when completing an AIA: The Office of the Chief Information Officer (OCIO) at the Treasury Board of Canada Secretariat (TBS) is responsible for maintaining the AIA tool and overseeing departmental compliance with the Directive on Automated Decision-Making.

Algorithmic Accountability Act EUA	Sim	Sec. 2 (12) IMPACT ASSESSMENT.—The term "impact assessment" means the ongoing study and evaluation of an automated decision system or augmented critical decision process and its impact on consumers.	Sim	To direct the Federal Trade Commission to require impact assessments of automated decision systems and augmented critical decision processes, and for other purposes.
Assembly Bill 331 on Automated Decision Tools (California)	Sim	22756.1. (a) On or before January 1, 2025, and annually thereafter, a deployer of an automated decision tool shall perform an impact assessment for any automated decision tool the deployer uses	Sim	Section 22756.7. (a) Within 60 days of completing an impact assessment required by this chapter, a deployer or a developer shall provide the impact assessment to the Civil Rights Department.
NIST	Sim	Risk management should be continuous, timely, and performed throughout the AI system lifecycle dimensions.	-	-
Committee on Artificial Intelligence - CAI	Sim	Art. 15 (2) (e) Ensure that the risk and impact management processes are carried out iteratively throughout the design, development, use and decommissioning of the artificial intelligence system;	Sim	Article 25 – Effective oversight mechanisms 1. Each Party shall establish or designate one or more effective mechanisms to oversee and supervise compliance with the obligations in the Convention, as given effect by the Parties in their domestic legal system. 2. Each Party shall ensure that such mechanisms exercise their duties independently and impartially and that they have the necessary powers, expertise and resources to effectively fulfill their tasks of overseeing compliance with the obligations in the Convention, as given effect by the Parties in their domestic legal system.

OECD	Yes	"AI actors should, based on their roles, the context, and their ability to act, apply a systematic risk management approach to each phase of the AI system lifecycle on a continuous basis to address risks related to AI systems, including privacy, digital security, safety and bias"	-	-
UNESCO	Yes	Member states and companies should implement adequate measures to monitor all phases of the life cycle of AI systems, including the functioning of algorithms used for decision-making, the data, as well as the AI actors involved in the process, especially in public services and where direct end-user interaction is required, as part of the ethical impact assessment.	-	These assessments should also be multidisciplinary, multi-stakeholder, multicultural, pluralistic, and inclusive. Public authorities should be obliged to monitor the AI systems implemented and/or used by these authorities, introducing appropriate mechanisms and tools.

Regulation	Mandatory Cases for Conducting Impact Assessments	Methodology	Implementation	Are there analysis criteria?
Bill 21/20	There's no mention of AIA, only a regulatory impact assessment	-	-	-
Bill 759/23	Whenever the system is considered high risk by the preliminary assessment.	Methodology based on risks and rights. Definition of at least 4 stages: preparation, risk awareness, mitigation of the risks encountered and monitoring.	Preliminary with the possibility/obligation of periodic reviews.	Yes, in Paragraph 1 of Article 24.

Canada's Artificial Intelligence and Data Act (AIDA)	For high-risk AI systems.	Risk-based methodology (not explicitly stated). It only determines the need for the identification, evaluation, and mitigation of risks of harm or biased outcomes beforehand.	Preliminarily.	-
Canada's Algorithmic Impact Assessment tool	For automated decision systems within the Public Administration.	Risk-based methodology. It only determines the need to identify risks and mitigate them. But it is available in an open questionnaire on the Government's Portal.	At the start of the projects' design phase and then before the system is produced to validate the results previously obtained.	Yes, there are tables with guiding questions for the identification of risk areas and for the identification of possible mitigation measures, and topic 3.2 contains points that should be considered when conducting the AIA.
Algorithmic Accountability Act EUA	For automated decision systems and extended critical decision processes (processes, procedures or other activities that use automated decisions to make critical decisions) of covered entities, in accordance with section 2(7).	It doesn't specify a precise methodology, but it seems to be risk-based.	Before and after system implementation.	Yes, as determined in Section 4.
Assembly Bill 331 on Automated Decision Tools (California)	Automated decision tools that meet the criteria of Section 22756.1.	Risk-based methodology	Annually, and as soon as possible after any significant update.	Yes, as determined in 22756.1. (a) (b).
EU AI Act (European Parliament version)	High-risk AI systems	Risk-based methodology	Prior to its use and at any other time when the deployer considers that there are new analysis criteria.	Yes, as determined in Article 29a (1).

CAHAI - Human Rights, Democracy and Rule of Law Impact Assessment of AI systems	No reference is made to an impact assessment, but there is mention of risk management and assessment	Risk-based methodology	-	-
CAI - Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (consolidated summary)	Does not define mandatory cases	Risk-based methodology (Article 15, 2)	It should be conducted iteratively throughout the design, development, use and decommissioning of the AI system.	Yes, as determined in Article 15 (2)
OECD[100]	There is no reference to an impact assessment, but there is mention of risk management and assessment.	Risk-based methodology	Before ("AI in the lab") and after ("AI in the field") its use/implementation.	Regardless of the number of risk levels, typical criteria for determining the level of an AI system include its scale (severity of adverse impacts (and likelihood), scope (breadth of application, such as the number of individuals who are or will be affected) and optionality (degree of choice as to whether to be subject to the effects of an AI system).
UNESCO - Recommendation on the Ethics of Artificial Intelligence + AI Ethical Impact Assessment Tool	Specific section on "ethical impact assessment," in which it defines that "Member States should create frameworks for conducting impact assessments, such as ethical impact assessment, to identify and assess the benefits, concerns, and risks of AI systems, as well as	Defines that Member States must adopt a normative framework that establishes a special procedure for public authorities.	Preferably before the technology is launched on the market but applied throughout its life cycle.	There are no specific criteria, but there is mention of "identifying impacts on human rights and fundamental freedoms, especially, but not limited to, the rights of marginalized and vulnerable people or people in vulnerable situations, labor rights, the environment and

100 https://www.oecd-ilibrary.org/docserver/2448f04b-en.pdf?expires=1691001152&id=id&accname=guest&checksum=4B452E3AB7BD695B35EF6D45563DC6B6; https://www.oecd-ilibrary.org/docserver/cb6d9eca-en.pdf?expires=1691002504&id=id&accname=guest&checksum=C786D4EDC16B1C3E6620F6617CCF0ADB.

	appropriate risk prevention, mitigation, and monitoring measures," but does not define in which cases there would be this obligation.			ecosystems and their ethical and social implications and facilitating citizen participation."

It should be noted, therefore, that one of the common threads in AI regulations is the provision for a minimum proceduralization of the algorithmic impact assessment tool. Unlike the UNESCO and Council of Europe guidelines, Bill 2338/23, as well as the European and Canadian proposals, is still timid when it comes to the component of possible adverse effects on social rights such as those related to work and the environment. Even so, only the Brazilian proposal is interoperable with all the others in the sense that it provides for and employs minimum normative density for it to flourish by including a methodological foundation, timing of analysis and possible review and criteria.

b.2) Transparency

In addition to the methodological definition, one of the essential aspects of impact assessments, especially when it comes to AI, is the definition of whether or not it is mandatory to disclose them. As mentioned above, Kaminski, in his studies on forms of risk regulation, approaches this regulation as linked to the idea of democratic supervision or democratic accountability[101]. In other words, risk assessment, within the process of constructing an impact assessment, would serve as an instrument for public discussion of these risks, which would then be shared with society as a whole.

In this context, the possibility of public access to the analysis or the main results of the algorithmic impact assessment process would allow the risk management of AI systems to be subject to public scrutiny, which would ensure that AI agents are accountable not only to the competent authorities for supervision, but to society as a whole, especially the individuals impacted by it, and could even serve as a potential basis for future substantive policy interventions (such as updating a possible list of high-risk AI systems).

According to UNESCO's Recommendation on the Ethics of Artificial Intelligence, (ethical) impact assessments should be transparent and open to the public, when deemed appropriate. In this sense, publicizing impact assessment analyses or their main conclusions would be in line with the idea of qualified transparency[102], facilitating the accountability processes for agents in the AI production chain and subsequent oversight by any competent authority and by the individuals and groups impacted by the system, including reducing asymmetries of information and power. Accordingly, all stakeholders can acquire the understanding and ability to influence the decision-making processes for creating and implementing AI systems, based on the idea of procedural justice and informational due process, which also lends greater legitimacy to the process[103].

The impact assessment algorithm can be publicized by making its entire content available, as well as its main conclusions[104]. In addition, the cases in which it is mandatory to publish such a tool may vary according to, for example, the degree of risk of a given AI system, the context of use or the type of AI agent (public or private sector). By way of example, in the field of data protection, both the GDPR and the LGPD do not define the

101 KAMISNKI, 2022, p. 36.

102 BIONI; LUCIANO, 2019, p. 3.

103 DARIUSZ, Kloza. *Privacy Impact Assessment as a Means to Achieve the Objectives of Procedural Justice*, Jusletter IT. Die Zeitschrift für IT und Recht, available at: https://cris.vub.be/files/49868387/Kloza_2014_PIA_as_a_Means_to_Achieve_the_Objectives_of_Procedural_Justice.pdf; CITRON, Danielle, PASQUALE, Frank. The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014.

104 For Bill 2338/2023, this disclosure can also take place in relation to the prior analysis process represented by the preliminary assessment in Article 13, with regard to AI systems within the public sector, regardless of the degree of risk (Article 21).

obligation to publish the results of the evaluation or impact report, respectively, which is only suggested in the first case and can be done as a good practice of accountability of the regulated agent in both the LGPD and the GDPR.

In the current regulatory proposals for AI worldwide, with regard to publicizing the AIA tool, Bill 5116 by the Washington State legislature stands out, as it seeks to establish criteria for the purchase and use of automated decision systems by the state. Unlike the GDPR and LGPD, section 5 expressly requires that that an algorithmic accountability report must be submitted for assessment by the competent body (called the "Office of Algorithmic Accountability Review" in the bill) on its official public website, as well as establishing a process for receiving comments from the public prior to its approval in a period of no less than 30 days[105].

Also in the United States, the Algorithmic Accountability Act of 2022, which targets automated decision-making systems, follows a different logic, since section 4(c) expressly states that it is not mandatory to disclose the impact assessment, despite the provision that a summary report of these assessments be sent, prior to implementation and on an ongoing basis, to the competent inspection and supervision body (section 3(b)(1)(d)(e)). Thus, despite the lack of public availability of the content or summary of the assessment to the general public, this regulatory initiative creates to a certain extent public accountability by mandating the sharing of the summary of the analysis to the competent body[106].

Similarly to the Washington bill, despite creating limits on industrial and commercial secrets, the Brazilian Bill 2338/2023 expressly provides for publicizing the conclusions of the impact assessment in its Article 26[107] determining a minimum content to be made available to the public, which includes, for example, a description of the purpose of the system, its context of use and territorial and temporal scope; risk mitigation measures adopted; and a description of the different segments affected. Also, considering the asymmetry of power in relations between the state and society, the bill provides, in item

105 Text of Section 5 (3): "(3) An agency intending to develop, procure, or use an automated decision system for implementation after January 1, 2024, must submit an algorithmic accountability report to the applicable algorithmic accountability review office and obtain approval or conditional approval prior to any use of the automated decision system. The algorithmic accountability review office must post the algorithmic accountability report on the algorithmic accountability review office's public website and invite public comment on the algorithmic accountability report for a period of no less than 30 days".

106 Kaminski (2022, p. 73 e 74) criticizes the lack of disclosure of the core of impact assessments, so that they remain internal to companies and are not disclosed to regulators, stakeholders, experts, or the public. In this "absence of public accountability," the author questions whether the regulated entities will actually mitigate the risks of their systems and remains skeptical about this.

107 Article 26 of Bill 2338/2023: "Article 26. Guaranteed industrial and commercial secrets, the conclusions of the impact assessment shall be public, containing at least the following information: I – a description of the intended purpose for which the system will be used, as well as its context of use and territorial and temporal scope; II – risk mitigation measures, as well as their residual level, once such measures have been implemented; and III – a description of the participation of different affected segments, if any, under the terms of § 3 of Article 24 of this Law.".

VI of Article 21, for extra measures for the public authorities with regard to publicizing preliminary assessments of AI systems developed, implemented, or used in this context, regardless of the degree of risk.

Along with Bill 2338/2023 (Brazil) and Bill 5116 (USA/Washington), the EU AI Act, in its June 2023 version published by the European Parliament, also incorporated this obligation to disclose the impact assessment (if concerning fundamental rights). In the European context, this rule is restricted to the summary of the impact assessment in cases of systems deployed by a public authority or certain organizations considered "gatekeepers" by Regulation (EU) 2022/1925 (Digital Markets Regulation) when they are considered deployers. Such information must be published in a public database of high-risk AI systems.

Bill 2338 also provides for the creation of a Brazilian database on artificial intelligence systems, which could contain both self-assessment documentation for AI systems and high-risk AI impact assessments. There are at least two objectives that the creation of the database fulfills. The first is transparency and the reduction of informational asymmetry in relation to those potentially impacted by AI systems. That is, those who have easier access to relevant information to assess risks to individual and collective rights and have knowledge about which AI systems affect their daily lives. The second is for the suppliers themselves, who can use the database to verify best practices in relation to the preparation of impact reports, promoting a culture of sharing and *benchmarking*[108].

The creation of a publicly accessible database with information on AI systems is not a Brazilian innovation, since it is also proposed, as previously mentioned, in the European Union's Artificial Intelligence Act , Chile's Bill 15869-19 and the United States' Algorithmic Accountability Act.

108 WRIGHT *et al.*, 2014, p.165.

Normativa	Previsão de publicização da AIA	O que deve ser publicado?	Onde	Previsão de banco de dados público	Onde
Bill 2338/2023	No	-	-	No	-
Bill 759/23	No	-	-	No	-
Bill 2338/23	Yes	For all high-risk AI systems: the main conclusions of the AIA; For public authority AI systems: all preliminary assessments, regardless of the degree of risk.	Article 26; Article 21	Yes	Article 43
Directive on Automated Decision-Making + Algorithmic Impact Assessment tool	Yes	For automated decision-making systems within the Public Administration: disclosure of the final results in an accessible format in English and French on the Open Government Portal.	Government of Canada website; Article 6.1.4 of the Directive	Yes	Government of Canada website; Article 6.1.4 of the Directive
Washington SB 5116 - 2021-22	Yes	-	Section 5 (3)	Yes	Section 6
Algorithmic Accountability Act EUA	Yes, in part	It only requires that a summary report of the evaluations be sent to the competent inspection and supervision commission.	Section 3 (b) (1) (d) (e) and section 4 (c)	Yes	Section 6
Assembly Bill 331 on Automated Decision Tools (California)	Yes, in part	It only orders the impact assessment to be sent to the Civil Rights Department.	Section 22756.7. (a)	No	-
EU AI Act (European Parliament version)	Yes	Only the release of the evaluation results summary in cases of systems deployed by a public authority or certain organizations considered “gatekeepers” by Regulation (EU) 2022/1925 (Digital Markets Regulation).	Article 29a (5)	Yes	Article 51 and 60

Committee on Artificial Intelligence - CAI[109]	Yes	Such measures shall take into account the risk-based approach referred to in Article 2 and: (g) require, where appropriate, publishing of the information about efforts to identify, assess, mitigate, and prevent risks and adverse impacts undertaken;	Article 15 (2) (g)	No	-
OECD[110]	Yes	Where appropriate (there are no further definitions), but it includes examples of what can be published: what governance mechanisms have been used, how risks are monitored and reviewed, what mechanisms exist for redress, among others.	Page 50 and 51 of the "Advancing accountability in AI" report	No	-
UNESCO[111]	Yes - impact assessments should be transparent and open to the public, where appropriate.	When appropriate (not defined).	Page 26, paragraph 53 of the Recommendations	Not mentioned	-
Blueprint for an AI Bill of Rights	Yes	Whenever possible, provided in a clear and machine-readable way, using simple language.	Page 5 and 28	Not mentioned	-

109 Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law – Committee on Artificial Intelligence (CAI). Strasburg, July 7, 2023.

110 OECD. Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI. 23 Feb. 2023. Available at: https://www.oecd-ilibrary.org/science-and-technology/advancing-accountability-in-ai_2448f04b-en.

111 UNESCO. Recommendation on the Ethics of Artificial Intelligence. 16 May 2023. Available at: https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.

b.3) Participatory democratic model

When it comes to impact assessments for certain technologies, services or products, statutory command can bring different levels of public participation and engagement. The involvement of interested parties can bring various benefits to the assessment process (for example, increasing its quality, credibility, and legitimacy) and to the outcome (for example, making the decision-making process better informed and representative)[112]. This means that decisions made based on impact assessments are not the result of an analysis restricted to a select group of stakeholders, especially internal to the organization, which could lead to bias and discrimination, but of diverse agents, including external ones, especially considering those potentially impacted by the implementation of the technology[113].

According to UNESCO, the preparation of the AIA requires the involvement of a range of potentially affected individuals, representatives, and communities, which can be done through multilateral consultations proportionate to the scale and scope of the system, its urgency and the expected impacts[114]. The broad and diverse participation of agents outside the regulated agent allows it to receive criticism and suggestions of possible impacts that were not thought of before the product/technology was launched, as well as enabling a transparent relationship between those impacted (present or future) and the agent, which is an instrument for reducing the information assymetry[115].

That said, it is essential that participation is effective, and to this end, preliminary documentation of the impact assessment should be made available to stakeholders so that they can conduct their own assessment[116]. That said, it is essential that participation is effective, and to this end, preliminary documentation of the impact assessment should be made available to stakeholders so that they can conduct their own assessment. This ensures that the process of preparing the impact assessment (and not just its outcome) is fair, which creates legitimacy, since people tend to trust decisions more when they are not taken behind closed doors, but involving people like them, as well as experts[117].

When it comes to AI, public participation in impact assessment processes is even more important in cases where difficult decisions have to be made, such as in the case of

112 KLOZA *et al*, 2019.

113 BIONI, Bruno; EILBERG, Daniela Dora; CUNHA, Brenda; SALIBA, Pedro; VERGILI, Gabriela. Data protection in the criminal and public security field: technical note on the Draft Data Protection Law for public security and criminal investigation. São Paulo: Data Privacy Brazil Research Association, 2020. p. 8–9.

114 UNESCO, 2023, p. 43.

115 WRIGHT *et al*., 2014, p. 160.

116 WRIGHT *et al*., 2014, p. 170.

117 ECNL; Society Inside. Framework for Meaningful Engagement. Disponível em: https://ecnl.org/sites/default/files/2023-03/Final%20Version%20FME%20with%20Copyright%20%282%29.pdf.

high-risk systems, since it potentially has relevant implications for fundamental rights, whether for marginalized groups or society in general[118]. The need for participation to be more than just a checklist item is even more eminent in this regard.

To this end, according to a study on Meaningful Public Engagement developed by the *European Center for Not-for-Profit Law Stichting*[119], the participatory process must consider three essential elements to be truly meaningful: (i) shared purpose, i.e. the purpose must go beyond the interest of the convening body itself and include the interests of those potentially affected or a general purpose in the public interest. (ii) credible process, i.e., inclusive, open, fair, and respectful, with minimal barriers of entry; and (iii) visible impact, in the sense that the involvement of the parties will have the power to contribute significantly to decision-making or introduce changes in the governance of the AI organization, product or service to align it with the public interest.

The more meaningful and inclusive, the more effective the stakeholder involvement is, both to understand potential problems or opportunities of products or services that use AI, and to identify possible specific impacts, implications, benefits, and harms, positive or adverse, of these products or services on individual and collective human rights. Especially considering the inclusion of marginalized and already vulnerable groups[120], which allows for the creation of systems that are better suited to the targeted social realities and with greater control by the population.

In order to better understand this dynamic, we can borrow the phrase "Nothing about Us without Us" to emphasize the importance of meaningful participation by society in the evaluation of AI tools, including to allow for democratic and social control of the agents[121]. This is in line with the principle of multi-stakeholder governance, which must be put into practice so that society is not just a passive agent of technology, but can act in its development, especially in cases of technologies that are supposed to have an impact on them.

According to Kaminski (2023, p. 79), because AI system risks are of varying degrees of unknown, unquantifiable, and socially contestable, the participation of different actors and stakeholders in the technology assessment process is a crucial aspect for the proper regulation of AI risk. Furthermore, this inclusion in the process of assessing the impacts of technology, based on the risk it has been classified as, also allows these groups to defend their rights more actively, which is in line with a regulation that follows a risk-

118 Ibid.

119 ECNL; Society Inside. Framework for Meaningful Engagement. Available at: https://ecnl.org/sites/default/files/2023-03/Final%20Version%20FME%20with%20Copyright%20%282%29.pdf.

120 Ibid.

121 COSTANZA-CHOCK, Sasha. Design Practices: "Nothing about Us without Us". Design Justice, published on 26 feb. 2020. Available at: https://designjustice.mitpress.mit.edu/pub/cfohnud7/release/4.

based and rights-based approach, and is an important tool for modern democracies[122].

In addition to technical experts, it is essential to involve those who are potentially most harmed by the application of certain AI systems, as they will be the ones who can describe the real risks and impacts related to the practical reality. This inclusion is also necessary to rebalance the unbalanced power dynamic between the organizations that build automated technologies and the people who use and are affected by them[123].

Consequently, for the scope of an AI impact assessment, it is indispensable to define stakeholders comprehensively, including agents internal or external to the AI agent, such as the public (laypeople), decision-makers, experts, civil society entities, academic researchers and all those who may be (today) or will be (in the future) impacted or impacting the AI system in question, especially vulnerable groups, and social minorities.

Moreover, for accountability purposes, it is also essential to record the participation of these stakeholders and the suggestions made for improving the AI system, allowing subsequent consultation to verify the effectiveness of participation and by other interested suppliers with possible similar impacts on their AI systems. When the impact assessment needs to be renewed, there should be public participation once again, albeit in a simplified way, depending on the level of change that has occurred between the initial consultation and the time of renewal.

As shown in the tables below, effective public participation throughout the process of preparing algorithmic impact assessments for AI systems, especially when they are high-risk, is a need advocated in different national laws, regulatory proposals, and suggestions from international organizations, in order to move towards the creation of impact assessment processes that are inclusive and permeable to public and citizen participation.

Internationally, the Council of Europe's CAHAI, in an evaluation study on impact assessment in Human Rights, Democracy and the Rule of Law, found that community engagement is essential to the success of this tool when applicable in the context of AI. To this end, it is important to define efficient mechanisms for identifying stakeholders within communities, in the most inclusive way possible, and, as a consequence, to produce meaningful active participation in system evaluation processes[124].

The need for diverse (internal and external) public participation, including ex-

122 BAROCAS, Solon; VECCHIONE, Briana; LEVY, Karen. Algorithmic Auditing and Social Justice: Lessons from the History of Audit Studies. EAAMO '21, October 5–9, 2021, –, NY, USA. Disponível em: https://dl.acm.org/doi/pdf/10.1145/3465416.3483294. p. 2.

123 Data & Society. Algorithmic Impact Methods Lab. Data & Society Announces the Launch of its Algorithmic Impact Methods Lab. Nova York, 10 mai. 2023. Disponível em: https://datasociety.net/algorithmic-impact-methods-lab.

124 CAHAI. Human Rights, Democracy and Rule of Law Impact Assessment of AI systems. Strasburgo, 11 mar. 2021. Conselho da Europa, CAHAI-PDG (2021)02. p. 15.

perts, civil society and affected communities (including those without technical knowledge), throughout the AI lifecycle, is also emphasized in other international documents, such as those from the OECD and UNESCO, as well as by US frameworks published by the *White House Office of Science and Technology Policy*[125] and by the *National Institute of Standards and Technology* (NIST)[126].

In the Brazilian scenario, with regard to bills 5051/19, 21/20, 871/21 and 579/23, due to their general nature, there is no provision for any instruments for democratic public participation throughout the lifecycle of AI systems. In Bill 2338/23, however, there is express mention of this qualified public participation at different times in the suggested text (§3 of Article 24, §2 of Article 25, Article 26, III and §2, point c, of Article 30). The possibility of democratic supervision, for example, is present in Article 26, by imposing the publication of the conclusions of the impact assessment, in §2 of Article 25, by determining public participation in updating the impact assessment, based on consultation with the parties.

In Brazil, this participation should be as inclusive as possible, including not only experts, but voices from different social groups, especially vulnerable groups (ranging from the Black population to traditional nations) so that cultural aspects, knowledge, and other distinctive characteristics are also taken into account in these evaluations, in order to avoid reinforcing the condition of under-representation, ethnic erasure and epistemicide[127].

The provision for democratic public participation in the impact assessment processes of AI systems also appears in draft regulations or regulations already in force from the European Union, the United States and Canada.

125 Blueprint for an AI Bill of Rights. Available at: https://www.whitehouse.gov/wp-content/uploads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

126 Artificial Intelligence Risk Management Framework (AI RMF 1.0). Available at: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.

127 Epistemicide is the term coined by Boaventura de Sousa Santos to explain processes of invisibilization and concealment of social and cultural contributions not assimilated by Western knowledge, as a result of colonial-capitalist structures and imperialist domination, especially of African and Indigenous peoples. SANTOS, Boaventura de Sousa. Constructing the Epistemologies of the South: Essential Anthology. Volume I: Towards an alternative thought of alternatives. Collection of Anthologies of Latin American and Caribbean Social Thought, 1st Ed, 2018., 1ª Ed, 2018.

Regulation	Provision for public participa–tion	In which terms?	Where?	Reinforced obligation for public author–ities?	Where?	External audit?	How?	Where?
Bill 21/20	No	-	-	-	-	No	-	-
Bill 759/23	No	-	-	-	-	No	-	-
Bill 2338/23	Yes	(i) No definition of those involved, mentions "differ-ent social seg-ments affected" and "interested parties"; (ii) It provides for this participation only at the time of updating	Article 24, § 3º Article 25, § 2º	Yes	Article 21, I and IV	Yes	It will fall upon the competent authority to regulate	Article 23, sole para-graph
GDPR UE	Yes	Where appropri-ate, the controller shall seek the views of data subjects or their representatives on the intend-ed processing, without prejudice to the protection of commercial or public interests of processing opera-tions	Article 35 (9)	Yes	Recital 93	Yes	Mentions the possibility of audits but does not regulate them	Article 28 (3) (h); Article 39 (1) (b); Article 47 (1) (j) and Article 58 (b);

California AB-2261 (2019-2020)	No	Before finalizing and implementing the "accountability report," the agency shall consider issues raised by the public through both of the following: (1) A public review and comment period as well as (2) Community consultation meetings during the public review period. This obligation is also present in the biennial revisions made to the accountability report	Section 1798.335. (e) (f) (g)	Yes	Section 1798.335.	Section 1798.365. (a)	Provision for external audit by the State Auditor of California.	Section 1798.370
Canada: Directive on Automated Decision-Making + Algorithmic Impact Assessment tool	Yes	It only mentions the need for consultation with internal and external stakeholders, including legal and privacy advisors; digital policy teams; and subject matter experts from other sectors.	The Government of Canada's website describes the tool	No	The directive applies to public authorities	Yes	It only mentions the possibility of external auditing, in cases approved by the Canadian government.	Section 6.2.5.2 and Section 6.2.5.3 of the Directive

Washington SB 5116 - 2021-22	Yes	It mentions the gathering of public comments + determines that the algorithmic accountability report must include a description of any public or community involvement conducted, as well as any future plans for public or community involvement in connection with the automated decision system.	Section 5 (3) and (6) (j) (v)	No	Project aimed at public authorities.	Sim	It provides for the possibility of audits by agencies.	Section 3(4) (b)
Algorithmic Accountability Act EUA	Yes	It determines the need for a meaningful consultation (including through participatory design, independent auditing or soliciting or incorporating feedback) with relevant internal stakeholders (such as employees, ethics teams and responsible technology teams)	Section 3, (b) (1) (g) and Section 4 (a) (2)	No	-	Yes	It mentions the possibility of an independent audit	Section 3, (b) (1)(g)

Algorithmic Accountability Act EUA		and independent external stakeholders (such as representatives and advocates of impacted groups, civil society and advocates, and technology experts) as often as necessary.						
EU AI Act (European Parliament version)	Yes	It provides for the involvement of representatives of people or groups who may be affected by the high-risk AI system, with the aim of receiving contributions to the impact assessment, with a period of six weeks for interested parties to respond. Small and medium-sized enterprises are exempt from this obligation.	Article 29a (4)	No	-	Yes	It provides for independent audits, especially in the case of analyzing compliance with the rules of the quality management system.	Recital 60-H, Art. 29 (5), Art. 70 (1) (b), Annex VII (5.3)

Council of Europe - CAI[128]	Yes	Incorporating the perspective of all stakeholders, including any persons who may have their rights potentially affected by the design, development, use or discontinuation of the AI system.	Article 15 (2) (c) and Article 19	No	-	No	-	-
OECD[129]	Yes	Consultation with stakeholders (internal and external), including civil society and affected communities (even without technical knowledge), to obtain feedback and knowledge to feed into impact and risk assessments, as well as managing that risk in each part of the process. Consultation should take	Page 52 of the "Advancing accountability in AI" report	No	-	Yes	In general, AI audits involve data scientists and engineers, models and systems, and governance experts, both internal and external. On this topic, they reinforce that the characteristics of the teams of auditors (such as gender, country, and	Mentions of auditing and the need for it are made throughout the "Advancing accountability in AI" report, especially on pages 24 and 47 et seq.

128 Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law – Committee on Artificial Intelligence (CAI). Strasbourg, July 7, 2023.

129 OECD. Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI. 23 Feb. 2023. Available at: https://www.oecd-ilibrary.org/scien-

		place at all stages of the AI system's life cycle. The format, cost and frequency of communications and consultations should be assessed based on the context.					ground) impact the evaluation of the fairness of the results of the AI system, which justifies the defense for the diversity of the teams that conduct these audits. Different levels of access could allow for audits and analysis adapted to a specific AI system and its context.	
UNESCO[130]	Yes	They must be transparent and open to the public when appropriate (does not define).	Page 26, paragraph 53.	No	-	No	-	-

ce-and-technology/advancing-accountability-in-ai_2448f04b-en.

130 UNESCO. Recommendation on the Ethics of Artificial Intelligence. 16 May 2023. Available at: https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence.

Blueprint for an AI Bill of Rights	Yes	"Automated systems should be developed with consultation from diverse communities, stakeholders, and domain experts to identify concerns, risks, and potential impacts of the system."	Page 15	No	-	Yes	There are no details, but it provides for independent evaluations to be conducted by third parties.	Mentioned a few times, such as on pages 20, 21, 24, 38 and 57.
Artificial Intelligence Risk Management Framework (AI RMF 1.0) - NIST	Yes	Identifying and managing the risks and potential impacts of AI requires a wide range of perspectives and stakeholders throughout its lifecycle. Ideally, AI stakeholders will represent a diversity of experiences, knowledge and backgrounds and comprise demographically and disciplinarily diverse teams.	Pages 9-10 and 29-31	No	-	Yes	It only mentions the possibility of an audit	Pages 16 and 35

		Experts, users, AI actors external to the team that developed or deployed the AI system and affected communities are consulted to support evaluations, as necessary.						
Executive Order on Safe, Secure and Trustworthy Development and Use of AI (EUA)	Yes	In terms of the use of AI by the federal government, it is now mandatory to implement minimum practices, based on the NIST framework or Blueprint, for managing AI risks that impact people's rights or safety, expressly mentioning that a public consultation must be held.	Seg. 10, 10.1, (b) (iv)	Yes	Seg. 10, 10.1, (b) (iv)	Yes	In the context of the use of automated decisions to implement social benefits, auditing is guaranteed.	Sec. 7, 7.2, (b) (ii) (E)

AXES 3 – Generative AI

As of November 2022, with the launch of ChatGPT by Open AI, generative AI quickly gained public attention, as demonstrated by the exponential growth of research and investment in this area since then[131]. Along with its many benefits, such as productivity and efficiency gains and helping to solve social challenges, there are also risks, which has led to an urgent discussion regarding its regulation. Defining generative AI is the first challenge in addressing the regulatory proposals. Just as there are multiple definitions of what Artificial Intelligence is, generative AI suffers from the same lack of a consensual definition.

In the terminology proposal signed between the United States and Europe, in the document *EU-U.S Terminology and Taxonomy for Artificial Intelligence*[132], which aims to "inform the approaches to AI risk management and Trustworthy AI on both sides of the Atlantic, and advance collaborative approaches in international standards bodies related to AI"[133], there is no definition of the term Generative AI, only of large language models (LLM):

> A class of language models that use deep-learning algorithms and are trained on extremely large textual datasets that can be multiple terabytes in size. LLMs can be classed into two types: generative or discriminatory. Generative LLMs are models that output text, such as the answer to a question or even authoring an essay on a specific topic. They are typically unsupervised or semi-supervised learning models that predict what the response is for a given task. Discriminatory LLMs are supervised learning models that usually focus on classifying text, such as determining whether a text was made by a human or AI[134].

Within the Chinese context, in July 2023, the Cyberspace Administration of China (CAC) published a few rules for generative AI in a document called "Interim Measures

131 OECD. G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI. Relatório preparado para a presidência japonesa de 2023 e para o grupo de trabalho digital e tecnológico do G7. Publicado em 7 set. 2023. Disponível em: https://www.oecd-ilibrary.org/deliver/bf3c0c60-en.pdf?itemId=%2Fcontent%2Fpublication%2Fbf3c0c60-en&mimeType=pdf.

132 European Commission. EU-U.S. Terminology and Taxonomy for Artificial Intelligence. Published on May 31, 2023. Available at: https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence.

133 Ibid, p.1.

134 Ibid, p.9. Original text.

for the Management of Generative Artificial Intelligence Services", which took effect on August 15th, 2023. Article 22 (1) of the document defines "generative artificial intelligence technology" as *"models and related technologies with the ability to generate text, pictures, audio, video and other content"*[135].

The existing bills in Brazil, on the other hand, do not provide definitions or rules for generative AI. Bill 2338/23, despite not providing a concept, mentions the term "general purpose artificial intelligence systems" which must include the indicated purposes or applications in their preliminary assessment (regarding their degree of risk) (Article 13, §1). As mentioned earlier in this report, the categorization of AI risks in Bill 2338/23 into excessive and high is based on the purposes and context of application.

In the European context, the first version of the AI Act proposal was published in April 2021, before the launch of the most famous LLM models, such as Open AI's GPT 3.5 and GPT 3. The first version of the proposal that initiated discussions on such models was the one adopted by the Council of the EU on December 6, 2022[136], which introduced the definition of "general purpose AI" in Article 3 (1b):

> 'general purpose AI system' means an AI system that - irrespective of how it is placed on the market or put into service, including as an open source software - is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others; a general purpose AI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems[137].

Recently, in the May 2023 version of the European Parliament (EP) document, amendment 169 defines general-purpose artificial intelligence systems as "systems that can be used in and adapted to a wide range of applications for which it was not intentionally and specifically designed"[138]. The EP's European proposal (2023) also introduced amendment 99, which provides a definition for *Foundation Models*:

135 http://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm. Originally in Chinese.

136 Council of the EU. Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights. Press Release, published on Dec. 6 2022. Available at: https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/.

137 Available at: https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf.

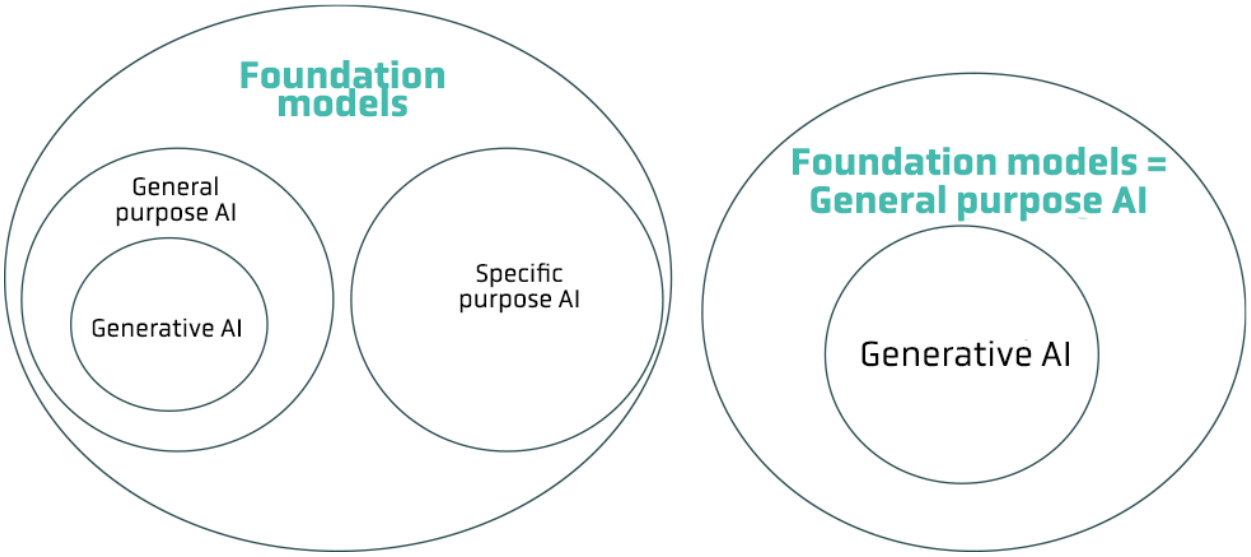
138 Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf. p. 113.

> Foundation models are a recent development, in which AI models are developed from algorithms designed to optimize for generality and versatility of output. Those models are often trained on a broad range of data sources and large amounts of data to accomplish a wide range of downstream tasks, including some for which they were not specifically developed and trained. The foundation model can be unimodal or multimodal, trained through various methods such as supervised learning or reinforced learning. AI systems with specific intended purpose or general purpose AI systems can be an implementation of a foundation model, which means that each foundation model can be reused in countless downstream AI or general purpose AI systems. These models hold growing importance to many downstream applications and systems[139].

As can be seen from the definitions adopted by Parliament, the relationship between the terms "general purpose AI" and "foundation or foundation AI model" is unclear, since, according to the definition, both specific purpose and general purpose AI systems can result from the use of a foundation model. Generative AI models, on the other hand, according to the European classification, are a type of foundation model: "(...) AI systems with the specific purpose of generating, with varying degrees of autonomy, content such as complex text, images, audio, or video" (amendment 399)[140]. In this case, one can understand that there is a relationship of genre and species, the genre being foundation models, and the species, generative AI models.

139 Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf. p. 74–75.

140 Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf, p. 200.



Possibilities for understanding the concepts introduced by the EP version of the EU AI Act

Based on European legislation, the definition of a foundation model could be adopted as encompassing both general-purpose AI and generative AI. Similarly, Hacker *et al*[141], addressing the regulation of AI models such as Chat-GPT, equates the terms "foundation models", "large language models" (LLMs) or "large generative AI models" (LGAIMs) - the latter chosen as the term adopted by this article.

From here on, alongside Hacker *et al*[142], we will refer to the designations used for AI models as comparable (generative AI, foundation models, large language models - LLMs, large generative AI models - LGAIMs), because, even if they don't designate exactly the same phenomenon, their common characteristics lead to similar considerations in relation to regulation, which is also followed by the OECD and the G7 countries[143].

More importantly, in addition to the challenge of conceptualizing regulation, which is common to the definition of AI itself, a second challenge arises from the fact that Generative AI puts a strain on risk-based regulation, the predominant model in attempts to regulate AI globally, as set out in this report. This is due to the fact that a regulatory model is eminently contextual, i.e., it depends on which specific situation the AI will be applied to in order to: a) assess possible risks to the rights of those impacted; and b) according to the risks, adjust the resulting obligations.

When it comes to LLMs, the models lend themselves to different purposes, which

141 HACKER *et al.*, 2023, p. 1113.

142 Ibid.

143 OECD. Initial policy considerations for generative artificial intelligence. Published on September 18, 2023. Available at: https://www.oecd-ilibrary.org/deliver/fae2d1e6-en.pdf?itemId=%2Fcontent%2Fpaper%2Ffae2d1e6-en&mimeType=pdf; OECD. G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI. Report prepared for the 2023 Japanese Presidency and the G7 Digital and Technology Working Group. Published on September 7, 2023. Available at: https://www.oecd-ilibrary.org/deliver/bf3c0c60-en.pdf?itemId=%2Fcontent%2Fpublication%2Fbf3c0c60-en&mimeType=pdf.

it might not have been possible to foresee when they were developed, thus challenging risk-based regulation, which focuses on regulating the uses of technology and addressing their impacts in specific contexts and from a less complex and dynamic chain of agents involved. Some provisions can be added in an attempt to mitigate or solve these challenges for the risk-based regulatory model.

First off, while not solving the challenge on its own, the provision for "general purpose artificial intelligence" within a risk-based regulation, establishing specific rules for such a model, is a first step towards its regulation. As mentioned, the version of the AI Act adopted by the EU Council on December 6, 2022, already included the conceptualization of such models, which was improved in the EP version of June 2023, which now also includes foundational and generative models. In the Brazilian context, most of the bills currently before the Brazilian Congress make no mention of these figures, with the exception of Bill 2338/23, which references "general purpose artificial intelligence systems" in §1 of Article 13, without, however, defining the term, as explained above.

In addition to introducing the idea of "general purpose AI" (and its variations), it is possible to think of regulating these AI models through risk-based regulation, by including within the idea of risk not only those that are known and predictable, but also risks that can foreseeably be expected, based on the principle of precaution. As such, even if there is no certainty as to the existence of certain risks, this uncertainty and lack of knowledge cannot be used as an excuse for not employing measures to prevent them from happening. This provision can be found both in the Brazilian context, in Bill 2338, and in the European context, in the latest version of Parliament's EU AI Act.

In Brazil, Bill 2338/23 establishes that if a general-purpose AI is used for one of the purposes listed as high-risk in Article 17, this system must comply with a series of governance obligations, including the preparation of an impact assessment, provided for through Articles 22 and 26. Within this assessment, the supplier of general-purpose AI must consider and record, among other elements, the "known and foreseeable risks associated

with the artificial intelligence system at the time, it was developed, as well as the risks that can reasonably be expected from it" (§1 of Article 24). To this end, the Brazilian project is in line with the precautionary principle, including providing that, in case of AI systems that could have irreversible or difficult-to-reverse impacts, the impact assessment should also consider incipient, incomplete, or speculative evidence (§2 of Article 24).

The Voluntary Code of Conduct on Responsible Development and Management of Advanced Generative AI Systems, announced in September 2023 by Canada's Minis-

ter of Innovation, Science, and Industry, also provides for "reasonably foreseeable risks" among the issues that should be analyzed by developers and managers in assessing the adverse impacts of generative AI systems to comply with the safety principle[144].

As for the European Union, the Council's version included a specific Title I(a) for general-purpose AI systems, which defines the rules applicable to providers of this technology in Article 4b, extending to them certain obligations for high-risk AI systems (despite mentioning the need for an implementing act specifying this application to general-purpose systems), without mentioning the "reasonably foreseeable risks". This term was added by the European Parliament's version in a new article also specifically created for foundation models. According to Article 28b (2), the supplier of these models must demonstrate, through appropriate design, testing and analysis, the identification, reduction, and mitigation of reasonably foreseeable risks to health, safety, fundamental rights, the environment and democracy and the rule of law before and during development. Accordingly, Section C of Annex III, which addresses transparency obligations for foundation systems, states that suppliers of these systems must make available and keep on record information on reasonably foreseeable risks and the measures that have been taken to mitigate them, as well as remaining unmitigated risks with an explanation as to why they cannot be mitigated.

Therefore, by introducing this idea of "reasonably foreseeable risks," what the Brazilian bill did, and which was later also included in the European regulation[145] and the Canadian code, was to try to ensure that, even if it is not possible to foresee all cases of risks associated with the foundation AI system, its suppliers should manage the risks that could reasonably be expected of them, even if they might not actually occur. This provision is in line with the precautionary principle.

Although the European proposal does not mention the precautionary principle, it can be drawn from Article 28b (2) and the new Recital 60-G. The latter establishes that, due to the complexity and unexpected impact of foundation AI systems, in addition to the lack of control of downstream AI providers over the development of LGAIMs, there must be a fair sharing of responsibilities along the AI value chain, which makes these

144 Government of Canada. Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. September 2023. Available at: https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems.

145 On June 14, 2023, the European Parliament adopted a series of amendments to the text of the EU AI Act. Among them is amendment 102, Among them is amendment 102, which introduced recital 60h, specific to foundation models; amendment 263 which added the term "reasonably" to the foreseeable risks in the risk identification and analysis stage of the risk management system; amendment 399 which introduced Article 28b to impose obligations of the provider of a foundation model; and amendment 771 which created Section C of Annex VIII to include reasonably foreseeable risks within the description of the capabilities and limitations of foundation models that need to be provided and recorded. For more information, please visit: https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf.

models subject to proportionate measures and more specific requirements and obligations, such as the obligation to assess and mitigate possible risks and harms and to implement data management measures, including the assessment of bias.

That said, the main difference between Bill 2338/23 and the EP version of the AI Act is in relation to the element that triggers the obligation to assess and mitigate risks for general purpose AIs. While the Brazilian model provides for the obligation to carry out an algorithmic impact assessment for the purposes of high-risk AI systems (Article 17), where LGAIMs may or may not fit in, the latest version of the European text immediately provides for specific obligations for these models in the new Article 28b, including the aforementioned assessment and mitigation of reasonably foreseeable risks, regardless of the level of risk.

For Hacker[146], the EP's version of the AI Act brings significant advances to the regulation of LGAIMs. However, according to the author's interpretation of Article 28b(2)(a), all these models would have to implement risk assessments and mitigation measures for reasonably foreseeable risks to health, safety, fundamental rights, the environment, democracy, and the rule of law, with the involvement of independent experts, which would make them, in practice, comparable to high-risk AIs. For the author, this "presumed" or "foreseeable" high-risk classification would make such models unfeasible in practice.

However, even if certain aspects of risk assessment are altered in the ideal regulation, considering their costliness under these models, it is important to point out that they already pose specific and relevant risks today, not only for human rights but also for the economy and society. These risks can be non-exhaustively divided into: (i) risks related to consumers; (ii) generation of disinformation; (iii) risks of restricting economic competition in the market; (iv) discrimination, (v) environmental sustainability; and (vi) artistic and intellectual property, especially with respect to copyright[147]. This alone would justify LGAIMs being subject to risk assessments and consequent mitigation.

Furthermore, Recital 60-G of the EP version of the AI Act states that specific requirements and obligations for foundation AI systems do not amount to considering these models as high-risk AI systems, but rather that their function is to ensure that there is a high level of protection of fundamental rights, health and safety, the environment, democracy, and the rule of law.

146 HACKER *et al.*, 2023, p. 1115.

147 An OECD study conducted in September 2023 analyzed some of these risks, which can already be experienced in practice, such as the amplification of disinformation, the reinforcement of discriminatory and biased practices, intellectual property rights issues and the impact on the labor market; OECD. Initial policy considerations for generative artificial intelligence. OECD Artificial Intelligence Papers. No.1. Published on September 18, 2023. Available at: https://www.oecd-ilibrary.org/deliver/fae2d1e6-en.pdf?itemId=%2Fcontent%2Fpaper%2Ffae2d1e6-en&mimeType=pdf.

In addition to the elements mentioned above, general-purpose AI systems can also be more effectively regulated using risk-based regulatory models by providing a better definition of the agents involved in the production chain of these systems, and by breaking down their obligations. A comparative analysis of Bill 2338/23 and the latest version of the AI Act from the European Parliament revealed a clear complexification of the network of agents involved in the European case. Whereas the Brazilian bill only mentions AI agents (supplier and operator), the European proposal includes suppliers, distributors, importers, operators and other third parties and, specifically for foundation models, includes the supplier, new suppliers and other agents involved in the systems' value chain.

A highlight of the EP's version was the proposal for cooperation between the players involved in LGAIMs, also included in the Council's version for suppliers (Article 4b(5)) but have been improved and made more complex in this latest update. According to Recital 60-F, for example, for foundation models provided as a service (such as those via API), it is stipulated as a rule that the original provider must cooperate with downstream providers throughout the period during which that service is provided and supported, in order to allow for appropriate risk mitigation. Moreover, the European Parliament's proposal, in Recital 60-G, addresses the element of uncertainty in the evolution of AI foundation models and how this impacts the definition of agents' responsibilities.

Thus, given the complexity of these models and the uncertainty surrounding them, the article highlights the need to clarify the role of the players who contribute to the development of these systems, especially the suppliers (original and subsequent). As the technology is complex and liable to cause unexpected impacts, especially as it can be used in a variety of ways, including for functionalities not initially thought of by the original supplier, the European text foresees a lack of control by subsequent suppliers and establishes stricter governance obligations for the original suppliers.

This is consistent with Recital 60-H, which states that, given the nature of foundation models, there is a lack of experience in compliance assessment and third-party auditing methods are still under development. Consequently, it establishes the obligation for the European Commission and the specific European AI authority, which will be created, to be responsible for periodically monitoring and evaluating the legislative and governance structure of such models in the EU context.

Although it does not provide a detailed breakdown of the chain of agents involved or include "reasonably foreseeable risks," due to its nature and scope[148], US President Joe

148 The Executive Order is significant but lacks the depth and detail that legislation can provide. While it has the power to initiate action and set priorities for federal agencies, it has no binding force on the business sector, despite its undeniable influence. This is because, by setting standards and requirements for the AI it acquires, the government can shape and direct market practices, since companies will have to adapt to these rules if they want to sign government contracts. In any case, the lack of enforceability and the presence of efficient governance mechanisms for companies

Biden's Executive Order on the "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence", published on October 30, 2023, addresses interesting points regarding the governance of foundation models by the federal government. It highlights the Biden administration's great concern about security risks related to the production of synthetic content, such as Generative AI, which can be negatively used to produce *deep-fake* e *deep nudes*[149], for example.

Among the measures announced are: (i) the promotion of capabilities for identifying and labeling synthetic content produced by AI, including the possibility of auditing such systems, the use of watermarking and a ban on generative AI producing child sexual abuse material or non-consensual intimate images of real people; (ii) obtaining information on dual-use foundation AIs whose model is fully open[150] (e.g. when the model is published on the Internet as they may pose significant safety risks, including soliciting input from the private sector, academia and civil society on the risks, benefits and policy and regulatory approaches applicable to these models including information on mechanisms for managing risks and benefits)[151].

Furthermore, the Executive Order has a specific section for promoting the effective and appropriate use of generative AI in the Federal Government (section 10, 10.1, (f)). After discouraging the imposition of broad general prohibitions or blocks on the use of generative AI by federal agencies, the document highlights, for example: (i) the need to limit access, as necessary, to specific generative AI services based on specific risk assessments; (ii) establishing guidelines and limitations on the appropriate use of generative AI; and (iii) encouraging the employment of risk management practices, such as employee training and compliance with recordkeeping, cybersecurity, confidentiality, privacy and data protection requirements.

And finally, on the subject of Generative AI, there is also discussion of the possibility of LLMs being audited. On this topic, Luciano Floridi et al (2023) advocates for an audit or risk assessment model for these models in three layers: governance, model, and application. According to the authors, LLM technology providers would first under-

---

can lead to non-uniform application and a lack of compliance; Center for AI and Digital Policy. "World Cup" of AI Policy News edition. CAIDP Update 5.42 – AI Policy News (Nov. 6, 2023). Available at: https://www.linkedin.com/posts/center-for-ai-and-digital-policy_caidp-update-542-ai-policy-news-nov-activity-7127339609293824000-f-CHi/.

149 Deep nude is the practice of using AI systems to generate fake nude content, usually based on a photo showing the victim dressed; LOPES, Larissa. Have you ever heard of Deep Nude? Jusbrasil, published in October 2023. Available at: https://www.jusbrasil.com.br/artigos/ja-ouviu-falar-na-pratica-do-deep-nude/1979706886.

150 The Biden administration's perspective seems to be in line with Irene Solaiman's theory, who advocates a framework for evaluating generative AIs according to their degree of openness/access fully closed; gradual or staged access; hosted access; cloud-based or API access; downloadable access; and fully open); SOLAIMAN, Irene. The Gradient of Generative AI Release: Methods and Considerations. February 2023. Available at: https://arxiv.org/abs/2302.04844.

151 See Section 4, 4.5 of the Executive Order.

go governance audits which would assess organizational procedures, internal accountability structures and quality management systems to verify, for example, the levels of robustness. Subsequently, the LLMs would undergo audits to assess their capabilities and limitations after initial training, but before implementation in specific concrete applications, in order to verify performance, information security and veracity. Finally, the products and services created based on the LLMs would undergo continuous application audits to assess legal compliance and their impact on users, groups, and the natural environment over time. These layers would act to inform and complement each other in order to contribute towards the proper governance of complex systems, including LLMs.

KEY PROPOSALS FOR REGULATING GENERATIVE AI IN BRAZIL				
Bill 5051/19	Bill 21/20	Bill 872/21	Bill 759/23	Bill 2338/23
No mention.	No mention.	No mention.	No mention.	Only mentions "general purpose artificial intelligence systems"

DRAFT BILLS THAT ALREADY MENTION GENERATIVE AI			
Regulation	Parameters		Mention in the regulation
Bill 2338/2023 (Brazil)	**Definition**	-	-
	Chain of involved agents	Provider and operator (AI agents)	Art. 4º, items II, III and IV and art. 13, § 1
	Provision of reasonable risks by the chain of agents	Yes	Art. 3, XI (general application) and Art. 24, § 1, (a) and § 2
	Obligations	There are no specific obligations for generative AI, but obligations for AI systems according to the degree of risk	Chapter IV (Governance of AI Systems)

EU AI Act Proposal - Council of Europe version	**Definition**	General purpose AI system	-
	Chain of involved agents	Provider	Article 4º, items II, III and IV and Article 13, § 1
	Provision of reasonable risks by the chain of agents	No	Art. 3, XI (general application) and Art. 24, § 1, (a) and § 2 (in case of high-risk AI)
	Obligations	The rules apply to high-risk AI systems but depend on an implementing act that would specify how these rules would apply to general-purpose AI systems, in light of their characteristics, technical feasibility, the specificities of the AI value chain and market and technological developments. There are exceptions to this rule.	Chapter IV (Governance of AI Systems)

EU AI Act Proposal - European Parliament version	**Definition**	Foundation model	-
		General-purpose artificial intelligence	Art. 3º, (1) (c)
		Generative AI	Art. 3º, (1) (d)
	Chain of involved agents	Operator (supplier, deployer, authorized representative, importer, and distributor). For generative AI, we're talking specifically about suppliers, new suppliers, and other actors in the AI value chain.	Art. 28b (4)
	Provision of reasonable risks by the chain of agents	Yes	Recital 60-H, Article 9 (2) (a), Article 28b (2) (a) and Annex VIII, Section c (6)
	Obligations	Different rules, regardless of whether it is provided as a stand-alone model or incorporated into an AI system or product, or provided under free and open source licenses, as a service, among others. Examples: (i) demonstrate the identification and mitigation of reasonably foreseeable risks, including the inclusion of experts in these assessments; (ii) incorporation of datasets only when subject to data governance measures; (iii) design and develop the foundation model, making use of applicable standards to reduce energy use; (iv) maintain technical documentation and intelligible instructions for use; (v) establishment of a quality management system; (vi) registration in the Art. 60 EU database; (vii) transparency obligations, among others.	Article 28b
***Interim Measures for the Management of Generative Artificial Intelligence Services* (China)**	**Definition**	Generative artificial intelligence technology	Article 22 (1)
	Chain of involved agents	Generative AI service providers and users	Article 22 (2) and (3) respectively

Interim Measures for the Management of Generative Artificial Intelligence Services **(China)**	Provision of reasonable risks by the chain of agents	No	-
	Obligations	There are various obligations for providers, such as protecting information, labeling content generated by generative AI, establishing complaints and denunciation mechanisms, as well as obligations for both providers and users, such as adherence to fundamental socialist values, measures to prevent discrimination, respect for intellectual property rights, transparency measures, etc.	Art. 4 and chapters 2 and 3.
Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems **(Canadá)**	Definition	-	-
	Chain of involved agents	Developers (differentiating between downstream developers) and managers	Table published on the official Code of Conduct website
	Provision of reasonable risks by the chain of agents	Yes	Measures to be taken in accordance with the Code of Conduct - Safety Principle
	Obligations	It defines a list of measures that should be taken according to principles, separating them into those that should be followed by deployers and managers, and varying whether it is a case of advanced generative systems for public use or not. Examples: implementation of a comprehensive risk management framework , information disclosure, cooperation between generative AI agents, testing methods, etc.	Table published on the official Code of Conduct website

***Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence - Joe Biden* (EUA)**	Definition	Dual-use foundation model and Generative AI	Section 3, (k) e (p)
	Chain of involved agents	-	-
	Provision of reasonable risks by the chain of agents	*Yes (there is no explicit term, but it could be from the concept of "AI red-teaming"*[152])	Section 3, (d)
	Obligations	Obligations for federal government agencies that include, for example, conducting a risk assessment for Generative AI (which can lead to limited access for a Generative AI), risk management practices, training personnel, creating a guide for the use of Generative AI in the workplace, conducting public consultations on widely available basic dual-use foundation models, among others.	Section 4, (i) (A) (B); Section 4, 4.4, (ii) (A) (B); Section 4, 4.5, (a), (iv); Section 4, 4.6; Section 8, (b), (i) (A); Section 10, 10.1, (b) (viii) (A), (B), (C); Section 10, 10.1, (f)

152 (d) "The term "AI red-teaming" means a structured testing effort to find flaws and vulnerabilities in an AI system, often in a controlled environment and in collaboration with developers of AI. Artificial Intelligence red-teaming is most often performed by dedicated "red teams" that adopt adversarial methods to identify flaws and vulnerabilities, such as harmful or discriminatory outputs from an AI system, unforeseen or undesirable system behaviors, limitations, or potential risks associated with the misuse of the system."

5. Brazilian Particularities for AI Regulation

When constructing a regulatory environment, it is essential that the regulation takes into account the particularities of the context in which it will be implemented. Brazil, as a country in the global[153], is crossed by social, racial, gender, colonial and territorial issues, which is reinforced by the widespread use of AI systems that carry within them mostly Western, European, white, and wealthy male voices and standards[154].

For example, the country still has serious human rights violations, leading the way in LGBTQphobia, inequality, racism, and gender violence[155]. Even today, we can speak of a digital colonialism, in which a large part of the majority world/global south, including Brazil, still finds itself in a position of colony, used to obtain cheap labor and extractive mining of data and raw materials, while also being placed in a condition of consumer market for emerging technologies from the global north, especially from large monopolistic technology companies[156].

Considering this current scenario of violence, inequalities and oppression, the country can't simply and unthinkingly import regulatory models from the global North, since we have our own circumstances (political, economic, and social), identities, characteristics, and problems - which must be taken into account when it comes to regulatory initiatives, thus demanding that the Brazilian regulation of artificial intelligence be thought of in terms of our peculiarities. Although we welcome the incorporation of foreign devices that make sense given our reality, it is crucial that specific devices are created to address our particularities and specificities.

In this context, Bill 21/20 and Bill 759/23 fail to make any progress on this issue, as they have generic wording, mostly principle based and with a low coercive load, since

153 The term Global South was first used in 1969 by political activist Carl Oglesby. The term was coined to replace expressions such as "underdeveloped countries" or "third world," which had negative connotations, as they reinforced stereotypes about poor communities and represented them as icons of poverty, masking their history of oppression and continuous exploitation. However, in recent times, the expression "global south" has also come to be seen as pejorative, since it ends up being inaccurate, homogenizing groups, as well as creating a certain geographical determinism, as if the countries of the Southern Hemisphere were fated to be poor and have no expectations of development. So perhaps the best expression is "Majority World," since these countries do in fact represent the majority of humanity; HEINE; Jorge. The Global South is on the rise – but what exactly is the Global South? National Interest, published on July 10, 2023. Available at: https://interessenacional.com.br/edicoes-posts/o-sul-global-esta-em-ascensao-mas-o-que-e-exatamente-o-sul-global/. Acesso em 25 jul. 2023; Demetriodor. O 'Sul Global' é um termo terrível. Não use! Publicado em 11 nov. 2018. Disponível em: re-design.dimiter.eu/?p=969; ARUN, Chinmayi. AI and the Global South: Designing for Other Worlds. In: DUBBER, M.; PASQUALE, F; DAS, S. Oxford Handbook of Ethics of AI. 2019.

154 AI Manyfesto, 2021.

155 SILVA, Tarcizio. Regulating artificial intelligence in Brazil could mitigate algorithmic racism. Folha de São Paulo, published on July 3, 2023. Available at: https://www1.folha.uol.com.br/blogs/politicas-e-justica/2023/05/regular-a-inteligencia-artificial-no-brasil-pode-mitigar-o-racismo-algoritmico.shtml#:~:text=Novo%20projeto%20de%20lei%20avan%C3%A7ou,combate%20aos%20danos%20do%20racismo&text=Os%20impressionantes%20saltos%20t%C3%A9cnicos%20nos,maravilha%20sobre%20as%20tecnologias%20digitais. Acesso em 21 jul. 2023.

156 FAUSTINO; LIPPOLD, 2023.

they do not provide for efficient governance tools to deal with the issues and risks of AI, especially in a scenario where these risks reinforce structural discrimination and are experienced in layers of intersectional oppression, as is the case in Brazil. By way of example, the term "non-discrimination" is only mentioned as a foundation and principle for the responsible use of AI in Brazil in Bill 21/20 and is not mentioned at all in the text of Bills 759, 872 and 5051.

More significant advances can be found in the text of Bill 2338/2023. Firstly, the bill recognizes the structural inequalities and asymmetries of the Brazilian context by expressly adopting the definitions of direct and indirect discrimination (Article 4, VI and VII) from the InterAmerican Convention against Racism, which, since 2022, has attained the status of a constitutional amendment in national territory, reinforcing protection against discrimination in different sections of its text[157].

One such moment occurs with the inclusion of a list of rights for individuals potentially affected by AI in Article 5, highlighting the rights to correct discriminatory biases (direct, indirect, illegal, or abusive - Article 12), to information (Article 7), to explanation (Article 8) and to contest (Article 9). In this respect, the bill reinforces that it does not prohibit the adoption of differentiation criteria when this occurs due to reasonable and legitimate objectives or justifications in light of fundamental rights (Article 12, sole paragraph), as is the case with affirmative action, such as racial quotas, for example.

In addition to focusing on the fight against discrimination, the text also addresses the need for protecting (hyper)vulnerable groups. To name a few examples, among the criteria for updating the list of high-risk and excessive-risk systems by the future AI authority is the fact that "the system is discriminatory" (Article 18, c) and that "the system affects people from a specific vulnerable group" (Article 18, d), and the methodology for the Algorithmic Impact Assessment highlights the possible discriminatory impact of the systems (Article 24, §1, f).

Furthermore, as previously addressed, Bill 2338/23 reinforces the importance of societal participation in assessing and understanding the risks of AI systems, by providing for this participation in Algorithmic Impact Assessment processes (Article 25, §2) and the obligation to publish its conclusions, which allows for public and social control of the risks.

As such, for the regulation of technology in the country, when compared to Bill 5051/19, Bill 21/20, Bill 871/21 and Bill 753/23, Bill 2338/23 can be interpreted not as a rival, but as a step forward, since its construction was thought out broadly and in uni-

157 Commission of Jurists Responsible for Supporting the Drafting of a Substitutive on Artificial Intelligence in Brazil (CJSUBIA). Final Report. 2022. Available at: https://www.stj.jus.br/sites/portalp/SiteAssets/documentos/noticias/Relato%CC%81rio%20final%20CJSUBIA.pdf p. 12–13.

son with society through public, multidisciplinary and multisectoral, national and international consultations and hearings, held throughout the CJSUBIA process in 2022, including greater inclusion of groups possibly affected by AI systems, especially minorities and vulnerable communities[158]. This made it possible to include and/or maintain regulatory aspects relevant to the Brazilian context, such as the definition of direct and indirect discrimination, the provision of special protection for vulnerable groups and the imposition of more robust governance measures when it comes to high-risk AI systems.

Bill 2338/2023 is more affirmative and protective of rights, which can be drawn from the broad list of foundations and principles set out in Articles 2 and 3, respectively. The broad list of principles in Article 3 also demonstrates the bill's concern with establishing a strong normative framework for protecting rights, taking into account not only internationally accepted AI principles such as reliability and robustness; transparency, explainability, intelligibility and auditability; accountability, responsibility and full reparation for damage; non-maleficence (to do no harm); and human participation, set out in documents from the OECD, European Union, Berkman Klein Center for Internet & Society, IEEE, G20 and public and private entities[159], but also others more applicable to the Brazilian reality, such as non-discrimination, justice, equity and inclusion; inclusive growth and sustainable development; due process of law, contestability and adversarial proceedings in a broad sense; prevention, precaution and mitigation of systemic risks.

This affirmative language of rights is in line with the constitutional values that stem from the 1988 Brazilian Federal Constitution (CF/1988), as well as other existing regulations that also prescribe a series of rights and safeguards, such as the General Data Protection Act (LGPD), the Consumer Protection Code (CDC), the Brazilian Civil Rights Framework for the Internet (MCI) and the Statutes of Racial Equality, the Elderly and Persons with Disabilities. Therefore, the conciliation of risk and rights regulation resonates through this semantic overture that connects the AI regulatory proposal to other legislation from which a dialog between all these normative sources[160]will necessarily emerge. This is without neglecting the constitutional concern of encouraging

158 Rights on the Net Coalition. Letter of Support for Bill 2338/2023. Published on June 14, 2023. Available at:: https://direitosnarede.org.br/2023/06/14/carta-de-apoio-ao-pl-2338-2023/. Accessed on July 18, 2023.

159 Fjeld, Jessica and Achten, Nele and Hilligoss, Hannah and Nagy, Adam and Srikumar, Madhulika, Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI (January 15, 2020). Berkman Klein Center Research Publication No. 2020-1, Available at SSRN: https://ssrn.com/abstract=3518482.

160 According to the Dialogue of Sources Theory, given the pluralism of legislative sources (international, supranational, and national), whether general or special, with converging fields of application, there needs to be dialogue and coordination between them, so that they are not revoked, derogated, or abrogated, but coordinated in favor of higher values, such as human rights and the protection of the vulnerable; MARQUES, Claudia Lima; BENHAMIN, Antônio Herman. The Dialogue of Sources Theory and its Impact in Brazil: a tribute to Erik Jayme. Revista de Direito do Consumidor (RDC) 2340, 115 indb, 2018. Available at: https://revistadedireitodoconsumidor.emnuvens.com.br/rdc/article/view/1042/911.

science, technology and innovation[161].

By way of example, Bill 2338/23 lists the need for the centrality of the human person, respect for human rights and democratic values and the free development of personality as the basis for the development, implementation, and use of AI in Brazil (Article 2), as expressed in articles 1 to 7 of the Brazilian Federal Constitution. These constitutional articles are also evidenced by the fundamental principles of defending equality, non-discrimination, plurality, and respect for labor rights; protecting privacy, data protection and informational self-determination; and preserving the environment and encouraging sustainable development. These fundamental principles, which provide for a more affirmative vocabulary of rights, are aligned with the encouragement of technological development and innovation, including through the promotion of research; and the defense of free enterprise, free competition, and consumer protection, which also radiate from constitutional values, especially from Articles 170, 205 and 218 of the Brazilian Federal Constitution.

The bill's concern with a balanced design of a normative structure that protects rights and its active role in promoting the country's economic development and innovation can also be seen in the topographical structure of the bill itself. This is due to the fact that the text first provides for a separate chapter to systematize the rights of those potentially impacted by AI and, at the end, a separate section with measures to foster innovation in the country.

Another striking aspect of Bill 2338/23 in terms of Brazilian particularities is the way in which it brings reinforced governance obligations to public authorities. If, in general, the relationship between state and citizen is naturally unequal due to the imbalance of power, including access to information, this imbalance of power is all the more intense in majority world countries like Brazil. The reason for this is that these are countries where the population still demands greater assistance and a welfare state is vital in the attempt to counter inequalities and achieve material equality.

For this reason, although there is no specific chapter for public authorities in Bill 2338, there are specific provisions for them. In section III of Article 14, for example, there is a possibility of an excessive AI system aimed particularly at the public authorities, which are prohibited from "evaluating, classifying or ranking natural persons, based on their social behavior or personality attributes, by means of universal scoring, for access to goods and services and public policies, in an illegitimate or disproportionate manner."

161 Speech by Professor Cláusula Lima Marques and Professor Danilo Doneda at the Parliamentary Session to install the Commission of Jurists (CJSUBIA) in the Federal Senate; TV SENADO. Artificial intelligence: setting up of the committee of jurists that will analyze the subject – 30/03/22. Held on March 30, 2022. Available at: https://www.youtube.com/watch?v=nXnliBi3vKY&ab_channel=TVSenado. Accessed on July 21, 2023.

The State has also been given special attention when listing the purposes for which AI systems are used, which are considered to be high-risk by Article 17, which includes, by way of example, "the assessment of criteria for access, eligibility, concession, review, reduction or revocation of private and public services that are considered essential", "administration of justice", "criminal investigation and public security", "investigation by administrative authorities" and "migration management and border control".

In addition, Article 21 creates additional governance measures for public bodies and entities to contract, develop or use high-risk AI systems. These include the requirement for public authorities to hold prior consultations and public hearings on the possible use of high-risk AI (item I), as well as having to guarantee, in a facilitated and effective manner, the right to human explanation and review to citizens in the case of decisions that generate relevant legal effects or significantly impact the interests of those affected (item IV)[162]. Another example of strengthening public transparency is the obligation to publish in easily accessible vehicles all the preliminary assessments of the AIs developed, implemented, or used by the public authorities, regardless of the degree of risk (item VI). These are measures that attempt to reduce the imbalance of power between the state and the citizen, especially in terms of knowledge and information, in order to ensure that AI reduces and does not amplify structural socio-economic distortions.

It is therefore important to emphasize that the current text of Bill 2338/23 represents an important first step for Brazil to regulate AI from a human-centric point of view, with the human person representing those who live and experience Brazil's structural asymmetries and inequalities (including racism). However, despite its irrefutable advances in terms of protecting rights, especially of vulnerable groups, and combating all forms of discrimination, there is still room for improvement.

Considering the deepening of inequalities and concentration of economic, political, and epistemic power in recent years as a result of the increased use of AI systems[163], Bill 2338/23 can advance in its anti-racist and anti-discriminatory commitment. For example, there are those who support the inclusion of the potential to reinforce the intersectional disparities present in the country as a criterion for evaluating excessive or high risk systems, as well as the express banning of AI systems[164] considered racist, sexist

162 GARROTE, Marina. Regulating Artificial Intelligence in Brazil. Center for Human Rights & Global Justice, NYU School of Law, published on September 28, 2023. Accessed in August 2023, but original article was taken down: https://chrgj.org/2023/09/28/regulating-artificial-intelligence-in-brazil/.

163 SILVA, Tarcizio. Regulating artificial intelligence in Brazil could mitigate algorithmic racism. Folha de São Paulo, published on July 3, 2023. Available at: https://www1.folha.uol.com.br/blogs/politicas-e-justica/2023/05/regular--a-inteligencia-artificial-no-brasil-pode-mitigar-o-racismo-algoritmico.shtml#:~:text=Novo%20projeto%20de%20lei%20avan%C3%A7ou,combate%20aos%20danos%20do%20racismo&text=Os%20impressionantes%20saltos%20t%C3%A9cnicos%20nos,maravilha%20sobre%20as%20tecnologias%20digitais. Acesso em 21 jul. 2023.

164 Ibid; Rights on the Net Coalition. Technical Note on Bill 2338/2023. August 2023. Available at: https://direitosna-

and transphobic[165], especially in sensitive contexts, such as facial recognition systems for public security or tools that evaluate the hazard posed by an individual for judicial purposes.

Furthermore, as advanced as Bill 2338/23 is in terms of protecting rights and seeking to counter discrimination, the text still takes a "defensive" stance, i.e., bringing in - necessary - governance instruments to promote defense against illegitimate or illegal results possibly produced by AI systems. However, the project has yet to make significant progress on "reactive" proposals, for example by encouraging the production of diverse, open, and multidisciplinary ethical AI databases and systems on national territory, combined with promoting education and training[166].

In this regard, Bill 2338/23 only mentions in item X of Article 2 "access to information and education, as well as awareness of artificial intelligence systems and their applications" as one of its legal foundations - while Bill 21/20 and Bill 759 fail to mention it at all. Brazil has good examples of legislation that sets out obligations for public authorities to train, raise awareness and educate in a concrete way, in accordance with constitutional values, as the Brazilian Civil Rights Framework (MCI) for the Internet did. This normative created a specific chapter for the actions of public authorities in favor of the inclusive development of the Internet in the country. To this end, for example, Article 26 of the MCI determines that, as a result of the State's constitutional duty to provide education, the population's qualification should include the safe, conscious, and responsible use of the internet as a tool for exercising citizenship, promoting culture and technological development.

---

rede.org.br/2023/08/23/coalizao-direitos-na-rede-divulga-nota-tecnica-sobre-o-pl-2338-2023-que-busca--regular-a-ia/.

165 BUOLAMWINI, Joy; GEBRU, Timnit. Gender Shades: intersectional accuracy disparities in commercial gender classification. Cambridge: Proceedings of Machine Learning Research, vol. 81, pp.1–15, 2018; BUOLAMWINI, Joy; RAJI, Inioluwa Deborah. Actionable Auditing: investigating the impact of publicly naming biased performance results of commercial AI products. Cambridge: Association for the Advancement of Artificial Intelligence/ACM conference on Artificial Intelligence, Ethics, and Society, 2019. Available at: https://www.media.mit.edu/publications/ actionable-auditing-investigating-the-impact-of-publicly-naming-biased-performance-results-of-commercial-ai-products/; COSTANZA-CHOCK, Sasha. Design Justice, A.I., and escape from the matrix of domination. Camrbidge: Journal of Design and Science, jul. 2018. DOI:10.21428/96c8d426. Available at: https://jods.mitpress.mit. edu/pub/costanza-chock/release/4; SILVA, Mariah Rafaela; VARON, Joana. Facial Recognition in the Public Sector and Trans Identities: technopolitics of control and the threat to gender diversity in its intersectionality of race, class, and territory. Research conducted by Coding Rights with support from the NGO Privacy International and funding from the International Development Research Center (IDRC). Rio de Janeiro: Jan. 2021; COSTA, Ramon; KREMER, Bianca. Artificial Intelligence and Discrimination: Challenges and Perspectives for the Protection of Vulnerable Groups in the Light of Facial Recognition Technologies. Fundamental Rights & Justice | Belo Horizonte, year 16, special issue, p. 145-167, October 2022. Available at: https://dfj.emnuvens.com.br/dfj/article/view/1316/1065.

166 SILVA, Tarcizio. Regular a inteligência artificial no Brasil pode mitigar o racismo algorítmico. Folha de São Paulo, published on July 3, 2023. Available at: https://www1.folha.uol.com.br/blogs/politicas-e-justica/2023/05/regular--a-inteligencia-artificial-no-brasil-pode-mitigar-o-racismo-algoritmico.shtml#:~:text=Novo%20projeto%20de%20lei%20avan%C3%A7ou,combate%20aos%20danos%20do%20racismo&text=Os%20impressionantes%20saltos%20t%C3%A9cnicos%20nos,maravilha%20sobre%20as%20tecnologias%20digitais. Acesso em 21 jul. 2023.

As far as regulating AI is concerned, there could be a programmatic chapter in Bill 2338/23 regarding the duties of public authorities, in collaboration with society. According to Professor Tarcízio Silva, *"Brazil has the human, historical and cultural wealth to lead the production of ethical digital technologies and combat knowledge bias in a multipolar world"*[167]. With state stimulus, which can come through programmatic norms when regulating AI, it is possible to invest in training the population in the use and development of AI systems for their safe, conscious, and responsible use.

Therefore, regardless of the possibilities for refinements in Bill 2338, the bill is now a less bumpy road to AI governance in line with Brazil's socio-economic context as a country located in the global south.

It is essential that Brazil builds an AI regulation that has similarities with foreign discussions and models, so that there is regulatory convergence, but taking into account the particularities of the country so that the regulation of technology works for the Brazilian context and for the people who live here, as was initiated through Bill 2338/2023. To paraphrase Cazuza, *"it's high time Brazil showed its face, or all that we'll be left with is the poor AI party to which we won't so much as be invited"*[168].

To that end, Bill 2338/2023 seems to be moving in the direction of building a more affirmative and protective vocabulary of rights[169], while also taking into account the country's particularities as a member of the majority world, as mentioned throughout this section. This approach is essential for the social advancement of unequal countries, as is the case in Brazil, so that AI and its regulation can serve the benefit of Brazilian society and not reinforce its harmful structural practices.

167 Ibid.

168 BIONI, Bruno; MENDES, Laura Schertel; ALMEIDA, Virgilio. Brazil could lead the way in regulating artificial intelligence. Folha de São Paulo, published on July 13, 2023. Available at: https://www1.folha.uol.com.br/ilustrissima/2023/07/brasil-pode-liderar-regulamentacao-da-inteligencia-artificial.shtml. Acesso em 18 jul. 2023.

169 Ibid.

REFERENCE TO BRAZILIAN PARTICULARITIES IN BILL 2338/2023	
Subject	Article contents
Affirmative rights language	A wide-ranging list of foundations and principles in Articles 2 and 3.
	Chapter II specifically provides for the rights of people affected by artificial intelligence systems.
Anti-discrimination and protection of vulnerable groups	Article 4 defines discrimination (VI) and indirect discrimination (IV).
	Section IV specifies the right to non-discrimination and to the correction of direct, indirect, unlawful, or abusive discriminatory bias - Art. 12: The sole paragraph of Art. 12 makes an exception for cases where differentiation criteria are adopted on the basis of demonstrated, reasonable and legitimate objectives or justifications in accordance with the right to equality and other fundamental rights.
	Art. 18 sets out the criteria for updating the list of excessive risk and high risk AI systems: the system has a high potential for material or moral harm, as well as being discriminatory (III) and the system affects people from a specific vulnerable group (IV).
	Art. 24, §1 stipulates that the impact assessment must consider and record at least the process and results of tests and evaluations and mitigation measures conducted to verify possible impacts on rights, with special emphasis on potential discriminatory impacts (f).
Proposals for reducing power asymmetries and strengthening public and social control of risks	Specific and additional governance obligations for the public authority for the development, contracting or use of high-risk AI systems provided for in art. 21.
	The possibility of society's participation in the assessment and knowledge of the risks of AI systems, based on the provision for this participation in the AIA (art. 25, §2).
	Obligation to publish the main conclusions of impact assessments in a high-risk artificial intelligence database, accessible to the public (Art. 43).

Concluding remarks

Regulatory interoperability: in between colonialism and normative emancipation

Currently, there is an effervescence of norms in which the discussion is increasingly not whether, but how to regulate artificial intelligence. Proposals are piling up at local, regional, and global levels of hard and soft laws that are not only difficult to follow, but above all to compare and understand their convergences and particularities.

This publication curated more than 20 (twenty) normative sources mapped along three (03) thematic axes: (i) risk-based regulation; (ii) algorithmic impact assessments; and (iii) Generative AI, ending with its own chapter on the national particularities of Brazilian AI regulation. Despite a common thread in terms of an asymmetrical and risk-based regulatory rationality (risk-based approach), we found that there is no homogenization, especially for the purposes of conciliation as an approach that is also affirmative of rights (rights-based approach). Some of these variations are mentioned below, repeating part of the conclusions drawn in the executive summary.

From a topographical point of view, there is variation in how the proposals organize not only concepts and principles, but mainly rights, with precedence given to the taxonomy of risks and good practices and governance measures. The choice to list rights first - preferably in their own chapter(s) - denounces that the ratio legis had as its primary, rather than secondary or even tertiary, point of attention the protection of people or groups affected by the benefits and risks of AI. Therefore, the normative structuring is also indicative of the much-desired reconciliation of a rights-based and risk-based approach.

The regulatory appetite is also significantly heterogeneous in listing which situations present unacceptable-excessive and high risk. From the difference between a ban and a moratorium on biometric data and artificial intelligence in the field of public security, to the length of the exemplary list of ex-ante bans on AI, to the quantitative and qualitative criteria for dynamic dilation-regression of the more or less intense regulatory burden to protect the affected individuals or groups. This interconnection between the logic of risk classification and rights can lead to the reinforcement or undermining of the implementation of the obligations at stake for the purposes of not just any kind of innovation, but one that is responsible.

A standout objective that reveals these nuances are the algorithmic impact assessments. If, on the one hand, this tool is unanimous, being listed in practically all the normative sources analyzed, on the other hand, the way it is dissected and minimally proceduralized is substantially different.

In this respect, Bill 2338/23 advances not only in terms of the need for a public version of such documentation - in line with the EU AI Act and other initiatives from Canada, the USA and Chile - but also, and above all, in the possible involvement of those impacted by the launch of the technology in a given context. A Brazilian regulatory legal tradition that undoubtedly stems from the environmental and consumer fields. There is a networked governance to set the law in motion. However, the proposal is still timid when it comes to the grammar of enhanced rights and impacts that go beyond the individual level, such as social impacts related to labor, the environment, culture, and others. This is an extremely important issue for the majority of countries in the world that suffer from the precarious extraction of labor (e.g., data labeling) and illegal mining (e.g., chip construction), which is a new type of colonization.

Thus, behind the movement towards regulatory interoperability, one should not minimize significant divergences with regard to the degree of "democratic oversight" of the tolerable risks associated with AI. This is a regulatory arrangement of greater or lesser public scrutiny whose distortions are historically evident in other regulatory experiences. For example, in the 2008 financial crisis, not only the regulatory system, but above all its enforcement , was captured, causing a systemic collapse in which social, economic, and technological development was structurally damaged for years[170]. In this tragedy, no responsible innovation took place and regulation was certainly one of these bottlenecks.

It is therefore urgent, especially for countries in the Global South, to see the convergences and, above all, the divergences of regulatory alternatives in terms of their degree of co-management over the risks of using AIs towards greater social porosity in order to unleash an emancipatory socio-technical approach[171]. It is necessary to remain vigilant in the game of so-called regulatory interoperability because there is a new type of colonialism at play. One that is more "insidious" and more "cunning"[172] in which rights and democratic oversight must not be emptied by the generic discursive narrative of asymmetrical risk-based regulation. Otherwise, effective accountability practices to re-

170 COHEN, Julie E. **Between Truth and Power: The Legal Constructions of Informational Capitalism**. Oxford University Press, 2019.

171 In this regard, for example, see the work developed by the Latin American Network of Studies on Surveillance, Technology and Society/LAVITS. Some notable works are: BRUNO, Fernanda. Algorithmic rationality & machinic subjectivity. IN: SANTAELLA, Lucia (Org.). Symbioses of the Human and Technologies: Impasses, Dilemmas, Challenges. São Paulo, SP: University of São Paulo Press /IEA-USP, 2022; BRUNO, Fernanda; PEREIRA, Paula Cardoso; FALTAY, Paulo. Artificial intelligence and health: resituating the problem. Electronic Journal of Communication, Information & Innovation in Health (RECIIS), vol. 17, nº 1, Apr-Jul 2023. Available at: https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/3842.

172 The terms in quotation marks and the idea defended are derived from SANTOS, Boaventura de Sousa. Boaventura de Sousa Santos: Colonialism and the 21st century. Fiocruz Center for Strategic Studies, published on April 6, 2018. Available at: https://cee.fiocruz.br/?q=boaventura-o-colonialismo-e-o-seculo-xxi.

duce informational asymmetry and, consequently, power[173], will not flourish.

173 According to Bruno Bioni: "What is at stake is not only the capacity for self-protection (...) but (...) how a plurality of actors will mobilize their respective prerogatives to reduce the asymmetry of power at stake. And thus experience a process of co-deliberation rather than informational domination;" BIONI, Bruno Ricardo. Regulation and Protection of Personal Data – The Principle of Accountability. São Paulo: Editora Forense, 2022. 320p. p. 245.

References

- Access Now. EU Trilogues: The AI Act must protect people's rights. Published on July 12, 2023. Available at: https://www.accessnow.org/press-release/eu-trilogues-ai-act/.
- Access Now. Joint statement: EU legislators must close dangerous loophole in AI Act. Published on September 7, 2023. Available at: https://www.accessnow.org/press-release/joint-statement-eu-legislators-must-close-dangerous-loophole-in-ai-act/.
- ACLU of Washington. How Automated Decision Systems are used in Policing. Published on 26 Dec. 2022. Available at: https://www.aclu-wa.org/story/how-automated-decision-systems-are-used-policing.
- AI Decolonial Manyfesto. Available at: https://manyfesto.ai/index.html.
- Artificial Intelligence Risk Management Framework (AI RMF 1.0). Available at: https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf.
- ARUN, Chinmayi. AI and the Global South: Designing for Other Worlds. In: DUBBER, M.; PASQUALE, F; DAS, S. Oxford Handbook of Ethics of AI. 2019.
- Data Privacy Brazil Research Association (DPBR). Technical Note - Data Privacy Brazil's contributions to Bill 21, of February 4, 2020. Available at: https://www.dataprivacybr.org/wp-content/uploads/2021/09/dpbr_notatecnica_pl21.pdf.
- BAROCAS, Solon; VECCHIONE, Briana; LEVY, Karen. Algorithmic Auditing and Social Justice: Lessons from the History of Audit Studies. EAAMO '21, October 5–9, 2021, –, NY, USA. Available at: https://dl.acm.org/doi/pdf/10.1145/3465416.3483294.
- BENNETT, Colin J.; RAAB, Charles D., Revisiting the governance of privacy: Contemporary policy instruments in global perspective. Regulation & Governance,Vol. 14, Issue 3, p. 447-464, 2018.
- BERNSTEIN, Peter L. Against the Gods: The Remarkable Story of Risk. Wiley, 1996.
- BIONI, B.; ZANATTA, R.; RIELLI, M. (2020). Data Privacy Br: Contribution to the Public Consultation on the Brazilian Artificial Intelligence Strategy. São Paulo: Reticências Creative Design Studio. Available at: https://www.dataprivacybr.org/wp-content/uploads/2020/06/E-BOOK-CONTRIBUIC%CC%A7A%CC%83O-DPBR-INTELIGE%CC%82NCIA-ARTIFICIAL-FINAL.pdf.
- BIONI, Bruno Ricardo. Regulation and Protection of Personal Data - The Principle of Accountability. São Paulo: Editora Forense, 2022, p.320.
- BIONI, Bruno; EILBERG, Daniela Dora; CUNHA, Brenda; SALIBA, Pedro; VERGILI, Gabriela. Data protection in the criminal and public security field: technical note on the Draft Data Protection Law for public security and criminal investigation. São Paulo: Data Privacy Brazil Research Association, 2020.
- BIONI, Bruno; LUCIANO, Maria. The Precautionary Principle for the Regulation of Artificial Intelligence: Would Data Protection Laws be its Gateway? In: Frazão, Ana. Mullholand, Caitlin. Artificial Intelligence and Law: ethics, regulation, and responsibility. São Paulo: Revista dos Tribunais, 2019.
- BIONI, Bruno; MENDES, Laura Schertel; ALMEIDA, Virgilio. Brazil could lead the way in regulating artificial intelligence. Folha de São Paulo, published on July 13, 2023. Available at: https://www1.folha.uol.com.br/ilustrissima/2023/07/brasil-pode-liderar-regulamentacao-da-inteligencia-artificial.shtml.
- BLACK, Julia. Proceduralisation and polycentric regulation. Revista Direito GV, Especial 1, pp. 099-130, 2005. p. 105-110.
- Blueprint for an AI Bill of Rights. Available at: https://www.whitehouse.gov/wp-content/up-

loads/2022/10/Blueprint-for-an-AI-Bill-of-Rights.pdf.

- BRADFORD, Anu. The Brussels Effect: How the European Union Rules the World. Nova York: Columbia Law School, mar. 2020.
- BRUNO, Fernanda; PEREIRA, Paula Cardoso; FALTAY, Paulo. Artificial intelligence and health: revisiting the problem. Revista Eletrônica de Comunicação, Informação & Inovação em Saúde (RECIIS), vol. 17, no. 1, Apr-Jul, 2023. Available at: https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/3842.
- BRUNO, Fernanda. Algorithmic rationality & machinic subjectivity. *IN*: SANTAELLA, Lucia (Org.). Symbioses of the Human and Technologies: Impasses, Dilemmas, Challenges. São Paulo, SP: Editora da Universidade de São Paulo/IEA-USP, 2022.
- BUOLAMWINI, Joy; GEBRU, Timnit. Gender Shades: intersectional accuracy disparities in commercial gender classification. Cambridge: Proceedings of Machine Learning Research, vol. 81, pp.1–15, 2018.
- BUOLAMWINI, Joy; RAJI, Inioluwa Deborah. Actionable Auditing: investigating the impact of publicly naming biased performance results of commercial AI products. Cambridge: Association for the Advancement of Artificial Intelligence/ACM conference on Artificial Intelligence, Ethics, and Society, 2019. Available at: https://www.media.mit.edu/publications/actionable-auditing-investigating-the-impact-of-publicly-naming-biased-performance-results-of-commercial-ai-products/.
- CAHAI. Human Rights, Democracy and Rule of Law Impact Assessment of AI systems. Conselho da Europa, CAHAI-PDG (2021)5. Strasbourg, May 21 2021. Available at: https://rm.coe.int/cahai-pdg-2021-05-2768-0229-3507-v-1/1680a291a3.
- CAHAI. Human Rights, Democracy and Rule of Law Impact Assessment of AI systems. Strasburgo, 11 mar. 2021. Conselho da Europa, CAHAI-PDG (2021)02.
- Center for AI and Digital Policy. "World Cup" of AI Policy News edition. CAIDP Update 5.42 - AI Policy News (Nov. 6, 2023). Available at: https://www.linkedin.com/posts/center-for-ai-and-digital-policy_caidp-update-542-ai-policy-news-nov-activity-7127339609293824000-fCHi/.
- CITRON, Danielle, PASQUALE, Frank. The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014.
- Rights on the Net Coalition. Letter of Support for Bill 2338/2023. Published on June 14, 2023. Available at: https://direitosnarede.org.br/2023/06/14/carta-de-apoio-ao-pl-2338-2023/. Accessed on July 18, 2023.
- Rights on the Net Coalition. Technical Note on Bill 2338/2023. August 2023. Available at: https://direitosnarede.org.br/2023/08/23/coalizao-direitos-na-rede-divulga-nota-tecnica-sobre-o-pl-2338-2023-que-busca-regular-a-ia/.
- Rights on the Net Coalition. Artificial Intelligence cannot be regulated at the drop of a hat. Published on September 23, 2021. Available at: https://direitosnarede.org.br/2021/09/23/inteligencia-artificial-nao-pode-ser-regulada-a-toque-de-caixa/; Rights on the Net Coalition. Brazil is not ready to regulate artificial intelligence. Published on December 7, 2023. Available at: https://direitosnarede.org.br/2021/12/07/brasil-nao-esta-pronto-para-regular-inteligencia-artificial/.
- COHEN, Julie E. Between Truth and Power: The Legal Constructions of Informational Capitalism. Oxford University Press, 2019.
- COLOMBO, Silvana. The mechanisms of popular participation in environmental management in the light of the constitutional text: positive and negative aspects. Editora Unijuí: Revista Direitos Humanos e Democracia, year 9, no. 18, jul/dec 2021.

- European Commission. EU-U.S. Terminology and Taxonomy for Artificial Intelligence. Published on May 31, 2023. Available at: https://digital-strategy.ec.europa.eu/en/library/eu-us-terminology-and-taxonomy-artificial-intelligence.
- Commission of Jurists Responsible for Supporting the Drafting of a Substitute on Artificial Intelligence in Brazil (CJSUBIA). Final Report. 2022. Available at: https://www.stj.jus.br/sites/portalp/SiteAssets/documentos/noticias/Relato%CC%81rio%20final%20CJSUBIA.pdf.
- National Council of Public Prosecutors. Collective Rights Portal. Available at: https://www.cnmp.mp.br/direitoscoletivos/.
- Consolidated Working Draft of the Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law - Committee on Artificial Intelligence (CAI). Strasburg, July 7 2023.
- COSTA, Ramon; KREMER, Bianca. Artificial Intelligence and Discrimination: Challenges and Perspectives for the Protection of Vulnerable Groups in the Light of Facial Recognition Technologies. Fundamental Rights & Justice | Belo Horizonte, year 16, special issue, p. 145-167, October 2022. Available at: https://dfj.emnuvens.com.br/dfj/article/view/1316/1065.
- COSTANZA-CHOCK, Sasha. Design Justice, A.I., and escape from the matrix of domination. Camrbidge: Journal of Design and Science, jul. 2018. DOI:10.21428/96c8d426. Available at: https://jods.mitpress.mit. edu/pub/costanza-chock/release/4;
- COSTANZA-CHOCK, Sasha. Design Practices: "Nothing about Us without Us". Design Justice, published on Feb 26 2020. Available at: https://designjustice.mitpress.mit.edu/pub/cfohnud7/release/4.
- Council of the EU. Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights. Press Release, published on Dec 6 2022. Available at: https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/
- DA SILVA, Paula Guedes Fernandes. Artificial Intelligence in the European Union: ways to regulate the technology that already regulates us. In: MENDES, Gilmar Ferreira; DE MORAIS, Carlos Blanco. Governance of the Changing Legal Order. Proceedings of the X Lisbon Legal Forum, 2022, p. 589. Available at: https://www.forumjuridicodelisboa.com/2023-anais.
- DA SILVA, Paula Guedes Fernandes; et al. Algorithmic impact assessment: what it is and how it is regulated in Brazil's Bill 2.338/23. Migalhas, published on Oct. 19, 2023. Available at: https://www.migalhas.com.br/coluna/migalhas-de-responsabilidade-civil/395547/avaliacao-de-impacto-algoritmico.
- DA SILVA, Paula Guedes Fernandes; GARROTE, Marina Gonçalves. Insufficiency of ethical principles to standardize Artificial Intelligence: anti-racism and anti-discrimination as vectors of AI regulation in Brazil. POLITICS, September 2022. Available at: https://politics.org.br/edicoes/insufici%C3%AAncia-dos-princ%C3%ADpios-%C3%A9ticos-para-normatiza%C3%A7%C3%A3o-da-intelig%C3%AAncia-artificial-o.
- DARIUSZ, Kloza. Privacy Impact Assessment as a Means to Achieve the Objectives of Procedural Justice, Jusletter IT. Die Zeitschrift für IT und Recht, available at: https://cris.vub.be/files/49868387/Kloza_2014_PIA_as_a_Means_to_Achieve_the_Objectives_of_Procedural_Justice.pdf.
- Data & Society. Algorithmic Impact Methods Lab. Data & Society Announces the Launch of its Algorithmic Impact Methods Lab. New York, May 10 2023. Available at: https://datasociety.net/algorithmic-impact-methods-lab.

- Data Privacy Brazil Research. Technical Note - Data Privacy Brazil's contributions to Bill 21 of February 4, 2020. Available at: https://www.dataprivacybr.org/wp-content/uploads/2021/09/dpbr_notatecnica_pl21.pdf.
- Data Privacy Brasil. Dadocracia - Ep. 78 - Legal Framework for AI. Dadocracia, published in Nov. 2021. Available at: https://open.spotify.com/episode/15BWzRa4cWVRo0jtGGPm4T?si=v7X-iVnWQ3eeIArlGmKaUg.
- Data Privacy Brasil. Dadocracia - Ep. 80 - Legal Framework for AI. Dadocracia, published in Dec. 2021. Available at: https://open.spotify.com/episode/0t4Rr07EwIjrdpmvzht79Z?si=OiOyUXc0T5-kH0nr6qhzKA&nd=1;
- Demetriodor. The 'Global South' is a terrible term. Don't use it! Published on Nov. 11, 2018. Available at: re-design.dimiter.eu/?p=969;
- ECNL; Society Inside. Framework for Meaningful Engagement. Available at: https://ecnl.org/sites/default/files/2023-03/Final%20Version%20FME%20with%20Copyright%20%282%29.pdf.
- Estadão. "Mais importante lei de recnologia no Brasil não está sendo debatida", diz especialista. Bruno Romani, published in Dec 7 2021. Available at: https://www.estadao.com.br/link/cultura-digital/mais-importante-lei-de-tecnologia-no-brasil-nao-esta-sendo-debatida-diz-especialista/.
- FAUSTINO, Deivison; LIPPOLD, Walter. Digital colonialism: towards a hacker-fanonian critique. 1st ed. São Paulo: Boitempo, 2023.
- FERRAZZO, Débora; DUARTE, Francisco Carlos. Legal colonization in Latin America. Available at: www.publicadireito.com.br/artigos/?cod=f376b8ae6217d18c.
- FIGUEIRA, Paulo Sérgio Sampaio. The role of the environmental council in public environmental policies. Published on April 14, 2022.Available at: https://direitoambiental.com/o-papel-do-conselho-do-meio-ambiente-nas-politicas-publicas-ambientais/.
- Folha de São Paulo. Brazil rushes through artificial intelligence law, say experts. Amanda Lemos, published on July 18, 2021. Available at: https://www1.folha.uol.com.br/mercado/2021/07/brasil-apressa-lei-para-inteligencia-artificial-dizem-especialistas.shtml.
- GAJARDONI, Fernando da Fonseca. Diffuse and Collective Rights I: General Theory of Collective Proceedings. São Paulo: Saraiva, 2012.
- GARROTE, Marina. Regulating Artificial Intelligence in Brazil. Center for Human Rights & Global Justice, NYU School of Law, published in Sep 28 2023. Available at: https://chrgj.org/2023/09/28/regulating-artificial-intelligence-in-brazil/.
- GASPAR; Walter B.; DE MENDONÇA, Yasmin Curzi. Artificial Intelligence in Brazil still needs a strategy. A Report by the Center for Technology and Society at FGV Direito Rio. May 2021. Available at: https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/30500/EBIA%20pt-br.pdf?sequence=3&isAllowed=y.
- GELLERT, Raphaël. Understanding the notion of risk in the General Data Protection Regulation. Computer Law & Security Review: The International Journal of Technology Law and Practice (2017).
- GOMES, Maria Cecília O. Between method and complexity: understanding the notion of risk in the LGPD. In: Current issues in data protection. PALHARES, Felipe (Coord.). São Paulo: Thomson Reuters Brazil, 2020, pp 245-271.
- GOMES, Maria Cecília. Data protection impact report: a brief analysis of its definition and role in the LGPD. AASP Magazine, n. 144, 2019. p. 10-11.
- Government of Canada. Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. September 2023. Available at: https://ised-isde.canada.

ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems.

- HACKER, Philipp. Sustainable AI Regulation. Privacy Law Scholars Conference 2023. Available at: https://arxiv.org/abs/2306.00292.
- HACKER, Philipp; ENGEL, Andreas; MAUER, Marco. Regulating ChatGPT and other Large Generative AI Models. Fairness, Accountability, and Transparency (FAccT '23), June 12–15, 2023. Available at: https://dl.acm.org/doi/10.1145/3593013.3594067.
- HEINE; Jorge. The Global South is on the rise - but what exactly is the Global South? National Interest, published on July 10, 2023. Available at: https://interessenacional.com.br/edicoes-posts/o-sul-global-esta-em-ascensao-mas-o-que-e-exatamente-o-sul-global/.
- HOOD, Christopher; ROTHSTEIN, Henry; BALDWIN, Robert.The Governance of Risk: Understanding Risk Regulation Regimes. New York: Oxford University Press, 2001. ISBN 0-19-924363-8.
- Jjeld, Jessica and Achten, Nele and Hilligoss, Hannah and Nagy, Adam and Srikumar, Madhulika, Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI (January 15, 2020). Berkman Klein Center Research Publication No. 2020-1, Available at SSRN: https://ssrn.com/abstract=3518482.
- KAMISNKI, 2022, p. 36; BOYD, Willian. Genealogies of Risk: Searching for Safety, 1930s-1970s. Ecology Law Quarterly, nº 895, 2012. Available at: https://scholar.law.colorado.edu/faculty-articles/143/.
- KlOZA, D., et al. (2017). Data protection impact assessments in the European Union: complementing the new legal framework towards a more robust protection of individuals. d.pia.lab Policy Brief, (1/2017), 1-4. https://doi.org/10.31228/osf.io/b68em, https://doi.org/10.5281/zenodo.5121575
- KLOZA, D., et al. (2019). Towards a method for data protection impact assessment: Making sense of GDPR requirements. d.pia.lab Policy Brief, 1(2019), 1-8. https://doi.org/10.31228/osf.io/es8bm, https://doi.org/10.5281/zenodo.5121534.
- LOPES, Larissa. Have you heard of the practice of Deep Nude? Jusbrasil, published in October 2023. Available at: https://www.jusbrasil.com.br/artigos/ja-ouviu-falar-na-pratica-do-deep-nude/1979706886
- MARQUES, Claudia Lima; BENHAMIN, Antônio Herman. The Dialogue of Sources Theory and its Impact in Brazil: a tribute to Erik Jayme. Consumer Law Journal (RDC) 2340, 115 indb, 2018. Available at: https://revistadedireitodoconsumidor.emnuvens.com.br/rdc/article/view/1042/911.
- Ministry of Science, Technology, and Innovations (MCTI). Brazilian Artificial Intelligence Strategy (EBIA). Available at: https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ebia-documento_referencia_4-979_2021.pdf.
- MONTELERO, A. (2022). Beyond Data - Human Rights, Ethical and Social Impact Assessment in AI. Asser Press, Information Technology and Law Series (IT&Law), Vol. 36.
- OECD. OECD Framework for the Classification of AI systems. OECD Digital Economy Papers, No. 323, OECD Publishing, Paris, 2022. Available at: https://doi.org/10.1787/cb6d9eca-en.
- OECD. Advancing accountability in AI: Governing and managing risks throughout the lifecycle for trustworthy AI. Published on Feb. 23, 2023. Available at: https://www.oecd-ilibrary.org/science-and-technology/advancing-accountability-in-ai_2448f04b-en.
- OECD. G7 Hiroshima Process on Generative Artificial Intelligence (AI): Towards a G7 Common Understanding on Generative AI. Report prepared for the 2023 Japanese presidency and the G7 digital and technological working group. Published on September 7, 2023. Available at: https://www.oecd-ilibrary.org/deliver/bf3c0c60-en.pdf?itemId=%2Fcontent%2Fpublication%2Fbf-

3c0c60-en&mimeType=pdf.

- OECD. Initial policy considerations for generative artificial intelligence. Published on September 18, 2023. Available at: https://www.oecd-ilibrary.org/deliver/fae2d1e6-en.pdf?itemId=%2Fcontent%2Fpaper%2Ffae2d1e6-en&mimeType=pdf.
- OECD. Common guideposts to promote interoperability in AI risk management. Published on Nov 07 2023. Available at: https://www.oecd-ilibrary.org/science-and-technology/common-guideposts-to-promote-interoperability-in-ai-risk-management_ba602d18-en.
- QUELLE, Claudia. 'The 'risk revolution' in EU data protection law: We can't have our cake and eat it, too' in R Leenes, R van Brakel, S Gutwirth and P De Hert (eds), Data Protection and Privacy: The Age of Intelligent Machines (Hart Publishing, forthcoming). 2017.
- QUELLE, Claudia. Does the risk-based approach to data protection conflict with the protection of fundamental rights on a conceptual level? Tilburg Law School Research Paper, 1-36, 2015.
- SANTOS, Boaventura de Sousa. Boaventura de Sousa Santos: Colonialism and the 21st century. Fiocruz Center for Strategic Studies, published on April 6, 2018. Available at:: https://cee.fiocruz.br/?q=boaventura-o-colonialismo-e-o-seculo-xxi.
- SANTOS Boaventura de Sousa. Building the Epistemologies of the South: Essential Anthology. Volume I: Towards an alternative thought of alternatives. Anthologies of Latin American and Caribbean Social Thought Collection, 1ª Ed, 2018.
- SILVA, Mariah Rafaela; VARON, Joana. Facial Recognition in the Public Sector and Trans Identities: technopolitics of control and the threat to gender diversity in its intersectionality of race, class, and territory. Research conducted by Coding Rights with support from the NGO Privacy International via funding from the International Development Research Center (IDRC). Rio de Janeiro: Jan. 2021;
- SILVA, Tarcizio. Timeline of Algorithmic Racism. Tarcizio Silva's blog, 2022. Available at: https://tarciziosilva.com.br/blog/posts/racismo-algoritmico-linha-do-tempo.
- SILVA, Tarcizio. Regulating artificial intelligence in Brazil could mitigate algorithmic racism. Folha de São Paulo, published on July 3, 2023. Available at:: https://www1.folha.uol.com.br/blogs/politicas-e-justica/2023/05/regular-a-inteligencia-artificial-no-brasil-pode-mitigar-o-racismo-algoritmico.shtml#:~:text=Novo%20projeto%20de%20lei%20avan%C3%A7ou,combate%20aos%20danos%20do%20racismo&text=Os%20impressionantes%20saltos%20t%C3%A9cnicos%20nos,maravilha%20sobre%20as%20tecnologias%20digitais.
- SOLAIMAN, Irene. The Gradient of Generative AI Release: Methods and Considerations. February 2023. Available at: https://arxiv.org/abs/2302.04844.
- Southern Alliance for the Global Digital Compact: contribution for the promotion of digital human rights. 2023. Available at: https://www.dataprivacybr.org/documentos/southern-alliance-for-the-global-digital-compact/.
- TRUBEK, David M.; COTRELL, Patrick; NANCE, Mark. "Soft Law," "Hard Law," and European Integration: Toward a Theory of Hybridity. Legal Studies Research Paper Series, Winsconsin, n. 1002, p. 1-42, nov. 2005. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=855447.
- TV Senado. Commission of jurists promotes debates on the regulation of artificial intelligence (2nd part) - 29/04/22. Published on April 29, 2022. Speech by Professor Maria Cecília Gomes. Available at: https://www.youtube.com/watch?v=P_yWp-2ZIZs&t=51s. Accessed on July 21, 2023.
- TV Senado. Artificial intelligence: establishment of the committee of jurists that will analyze the subject - 30/03/22. Held on March 30, 2022. Available at: https://www.youtube.com/watch?v=nX-nliBi3vKY&ab_channel=TVSenado.

- UNESCO. Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence. Published in 2023. Available at: https://unesdoc.unesco.org/ark:/48223/pf0000386276/PDF/386276eng.pdf.multi.
- UNESCO. Recommendation on the Ethics of Artificial Intelligence. Adopted on November 23, 2021, and published in 2022. Available at: https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence;
- VARON, Joana; SILVA, Mariah Rafaela. Facial recognition in the public sector and trans identities: technopolitics of control and threats to gender diversity in its intersectionality of race, class, and territory. Available at: <https://codingrights.org/docs/rec-facial-id-trans.pdf>.
- WRIGHT, David et al. Integrating privacy impact assessment in risk management. International Data Privacy Law, v. 4, n. 2, p. 155–170, 2014.
- YANG, Zeyi. China just announced a new social credit law. Here's what it means. MIT Technology Review, published on Nov. 22, 2022. Available at: https://www.technologyreview.com/2022/11/22/1063605/china-announced-a-new-social-credit-law-what-does-it-mean/.
- ZANATTA, Rafael A. F. Personal Data Protection as Risk Regulation: a new technical framework? 1ST MEETING OF THE INTERNET GOVERNANCE RESEARCH NETWORK, NOVEMBER 2017. Available at: https://www.redegovernanca.net.br/public/conferences/1/anais/ZANATTA,%20Rafael_2017.pdf.