ETHICAL IMPLEMENTATION OF AI IN JOB CANDIDATE RECRUITMENT: INSIGHTS ON DATA PROTECTION, AI, AND LEGAL PERSPECTIVES

Jessica Fernandes Rocha *

Abstract: This work aims to explore how workers, as data subjects, experience the effects of AI in recruitment, facing multiple layers of vulnerability as conceptualized by Florencia Luna in her theory of 'Layers of Vulnerability'. Following this reasoning, it examines the scope and limitations of personal data protection laws, with a particular focus on Law No. 13,709/2018 (LGPD), which aims to safeguard individuals' rights within this context. To envision future developments, the discussion will include legislative initiatives on the subject, emphasizing the importance of participation from all social actors in promoting the responsible use of AI in employment relations.

Keywords: personal data protection; Artificial Intelligence; employee recruitment.

INTRODUCTION

The influence of technological evolution and globalization has led to increasing demands in a labor market that has become increasingly competitive, motivating companies to optimize all resources, including human resources, within the productive environment. This effort has driven a transformation in the mindset of corporations aimed at maintaining competitive advantages. Pontes teaches that "people have always been important in the life of organizations; however, in the era of intellectual capital, individuals have become the vital factor for the maintenance of companies' competitiveness."¹ Attracting and retaining talent has, thus, become a vital element for the continuity of business.

To enhance the accuracy of hiring, inherent in the processing of curricular information about numerous candidates, Artificial Intelligence ("AI") tools with profiling techniques are employed. According to Hildebrandt, their objective is to "individuate and represent a subject or to identify a subject as a member of a group or category."² In practical terms, this functionality is

---

² Hildebrandt defines "profiling" as “The process of ‘discovering’ correlations between data in databases that can be used to identify and represent a human or nonhuman subject
especially relevant for recruitment issues because the goal is not to find an individual of singular identity but an individual who possesses a specific professional profile suitable for a particular position. On the other hand, the recruitment of job candidates using AI tools, often obscure and biased, exacerbates the existing power asymmetry in the employer-employee relationship, especially considering that the user-developer relationship is also marked by imbalances. Thus, individuals in the workforce experience the cumulative effects of these relationships, detrimentally affecting their legal status as workers, data subjects, and human beings. An aggravating factor to this scenario, as Benjamin Ruha teaches, is the fact that these discriminatory and opaque algorithms now appear as the new contractors and supervisors of the working class. At the intersection of privacy rights, labor rights, and fundamental rights, it is necessary to systematically interpret the legal framework, disputes, and perspectives in the context of recruitment through automated decisions. It is essential to consider that the scope of protection provided by personal data protection laws in these contexts has limitations, but there are initiatives to regulate artificial intelligence in Brazil and worldwide. Certainly, these initiatives represent an important element in ensuring that "this technological tool is led by humans to protect human values."

I. LAYERS OF VULNERABILITY OF WORKERS AS DATA SUBJECTS

Florence Luna proposes an approach to vulnerability analysis centered on layers, contrasting sharply with the idea of labels. This conception is built on the notion that "layers" adequately convey two highly important traits of this condition: i) the highly contextual nature of vulnerability, suggesting that

(individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category”. For a more in-depth conceptualization of profiling, see: Hildebrandt, M. "Defining Profiling: A New Type of Knowledge?" In Profiling the European Citizen. Cross-Disciplinary Perspectives, edited by M. Hildebrandt & S. Gutwirth, 17-30. Dordrecht: Springer, 2008. Available at: https://doi.org/10.1007/978-1-4020-6914-7_2

3 It is not the aim of this work to detail Artificial Intelligence (AI). For a more in-depth conceptualization of AI, see: Parentoni, L. "What Should We Reasonably Expect From Artificial Intelligence?" ResearchGate, 2022. Available at: https://www.researchgate.net/publication/361988480_What_should_we_reasonably_expect_from_artificial_intelligence.


attributing it to specific subpopulations offers a simplistic solution to a complex problem; and ii) the possibility that multiple layers of vulnerability could affect the same individual, causing them to experience its effects in unique ways.

Luna's theory is of utmost relevance because point i) accurately represents the fact that certain groups may be even more vulnerable in contexts of algorithmic discrimination, depending on the criteria programmed into the AI and/or the dataset used in its training. Despite the known reduction in opportunities for certain groups, some social markers may be particularly targeted by algorithmic biases, exacerbated by the AI's interaction with data reflecting current, exclusionary, and non-representative power structures.

Furthermore, highlighting the relevance of point ii), it is evident that regardless of the social markers affecting both the employer and the job seeker, the context of needing to establish oneself in the job market for subsistence adds a layer of vulnerability to the worker, compromising their ability to consent. The submission of an individual's data to the automated decisions of an AI, whose operations are unknown to them, introduces another layer.

One of the main, if not the primary, aggravating factors in the analyzed context lies in the fact that this hindrance to the rights of the personal data subject occurs within an inherently unbalanced relationship: the employment relationship. Pratyusha Ria Kalluri provides a highly enriching perspective for this analysis, emphasizing that one should not question whether an AI is "good" or "fair," but rather how it shifts power. Thus, any conclusions regarding how recruitment technologies change power must acknowledge the starting point of asymmetry.

Therefore, faced with the inherent hypo-sufficiency of the working individual, it is crucial to prevent this technology from intensifying the

---

6 The author does not deny the fact that some people are subject to greater risks than others but focuses on the point that some contexts are especially responsible for introducing more or fewer layers of vulnerabilities into certain individuals. Luna, F. "Elucidating the Concept of Vulnerability: Layers Not Labels." *International Journal of Feminist Approaches to Bioethics* 2, no. 1 (2009): 121–139. Available at: http://www.jstor.org/stable/40339200, 123.

7 An algorithm programmed to be discriminatory is one that was intentionally developed to unlawfully discriminate based on certain attributes (features), such as race, gender, or ethnic origin. However, there is a possibility that the algorithm may become biased due to machine learning processes. In this case, although it was not necessarily programmed to be discriminatory, it tends to replicate exclusionary and non-representative structures present in the database used for its training, seeking to maintain coherence between input data (inputs) and output data (output).

contextual vulnerability\(^9\) already affecting this individual. This principle, acting as a counterbalance to the worker's layer of vulnerability, is also applicable to job applicants, i.e., individuals who have not yet entered into a formally established legal employment relationship but are attempting to enter the job market. Paulo Gustavo de Amarante Merçon emphasizes that "the disadvantage in the legal relationship will affect the worker in various ways, from the constraint of their will to the vulnerable position in contract negotiation and the demand for its fulfillment."\(^10\)

In the case at hand, there appears to be an accumulation of vulnerability, which is one of the central proposals of this paper. Hence, each of the spheres composing this intersection will be further elucidated in sections 2.1 and 2.2. Building upon Kalluri's assertion that analyzing the malevolence or benevolence of an AI is futile, the intention is to highlight that the focus of this work lies in how these technologies impact the analyzed power relationship. In other words, in a relationship already marked by the worker's disadvantage, there is an even greater risk to their legal assets when transparency regarding (automated) decisions that directly affect their life and opportunities is completely removed.\(^11\)

A. Power relations between employees and employers

When analyzing the concept of "hypo-sufficiency," which permeates the entire sphere of labor law, it is evident that it is rooted in discussions about power, more specifically, the power disparity between parties. In other words, it describes an asymmetric relationship where the involved parties do not start from an equitable position, rendering a completely free dialogue impossible.\(^12\) This is because one party's subsistence is inherently tied to their entry into—and maintenance within—the job market, while the other party has the flexibility to replace them from a broad pool of available workers, justifying the Principle of Protection. Maurício Godinho Delgado elucidates


\(^10\) Merçon, P. G. de A. "Direito do Trabalho Novo." Revista do Tribunal Regional do Trabalho da 3ª Região 51, no. 81 (2010), 139.

\(^11\) These damages to the individual, referring to the reduction of their life opportunities as a result of the illicit processing of their data, are in direct conflict with the constitutional and legal principles provided for in Article 6 of the LGPD.

\(^12\) It is noted how the context in question is not compatible with the requirements for valid consent brought by the Brazilian General Data Protection Law in its Article 5, XII. Brazil. Lei n. 13.709, de 14 de agosto de 2018, Lei Geral de Proteção de Dados Pessoais (LGPD). Available at: http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2018/Lei/L13709.htm.
that this concept aims to "rectify (or mitigate) the legal imbalance that is inherent in the factual plane of the employment contract."\textsuperscript{13}

Beyond doctrinal positions in labor law, the relationship between vulnerability and consent is recognized in various fields of study. For example, in the realm of research ethics, the Belmont Report\textsuperscript{14} highlights this factor among three requirements for the concept of vulnerability: impairment of the capacity to consent. Florencia Luna critiques models of rational choice, highlighting that they "assume idealized abilities to choose rationally, while some accounts of justice overlook vulnerabilities that arise from \textit{subordination}\textsuperscript{15} and \textit{dependence} on others\textsuperscript{16} (emphasis added).

The CIOMS Guidelines aim to offer a general definition of vulnerability, describing it as some individuals' inability to protect their own interests, with the lack of power being the primary criterion. The commentary on Guideline 15, "Research Involving Vulnerable Persons and Groups," sheds light on factors that can render individuals vulnerable, such as "relative or absolute impairments in decisional capacity, education, resources, strength, or other attributes needed to protect their own interests."\textsuperscript{17}

Returning to the concept of vulnerability addressed in section 1.3, it becomes evident in this context that one party is highly susceptible to harm, positioned imbalancedly considering power dynamics, while the other party is in a considerably safer position. These contrasts involving "power relations," "vulnerability," and "capacity to consent," which are deeply interconnected across various fields of study, when examined in the context of data protection in employment relationships, challenge fundamental principles of privacy laws such as "Free Consent"\textsuperscript{18} and "Informational Self-

\textsuperscript{15} This concept is connected to the dependence inherent to the employment relationship provided for in Article 3 of the CLT, which provides “Any natural person who provides services of a non-occasional nature to the employer, under the latter's dependence and for a salary, is considered to be an employee”, freely translated.
\textsuperscript{18} The LGPD provides for consent in a highly qualified manner, conceptualizing it in its Article 5, Section XII, as “free, informed and unambiguous manifestation whereby the data subject agrees to her/his processing of personal data for a given purpose”. Translated by IAPP (International Association of Privacy Professionals), available at https://iapp.org/resources/article/brazilian-data-protection-law-lgpd-english-translation/
Determination.\footnote{Provided as one of the foundations of the LGPD in its Article 2, Section II.}

B. Power relations between users and developers

Similar to the dynamics discussed in the previous section, an asymmetry is observed in the relationship between AI developers and individuals whose data is subjected to AI processing. Highlighting the parallels between the scenarios of Users-Developers and Employees-Employers, scholars emphasize that this disparity impedes free dialogue in both contexts.

Sunstein notes that the vast majority of individuals lack the capacity to comprehend the workings of algorithms, suggesting that even a basic explanation of their functionality could significantly enhance trust levels.\footnote{Sunstein, C. R. "The Use of Algorithms in Society." SSRN, December 2022. Accessed November 20, 2023. https://ssrn.com/abstract=4310137.} He further mentions that the importance of this issue has led to the emergence of organizations like Data & Society, whose primary aim is to bridge the gap between technology creators and users by identifying biases, issues, and the impacts of technology adoption in daily life—a promising initiative.

Schermer points out that in cases of profiling, Informational Asymmetry persists, as the position of the data controller (in terms of information) improves, while that of the data subject does not change. According to Schermer, informational asymmetries contribute to "disturbing the current balance of power between different parties."\footnote{Schermer, B. "Risks of Profiling and the Limits of Data Protection Law." In Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases, edited by Bart Custers et al. Heidelberg: Springer, 2013, 139.} He also identifies this as a specific challenge occurring in two main scenarios: i) when the data subject is unaware of the profiling activity, or ii) when the data subject has incomplete information about the profiling process.

Therefore, one of the primary rights of data subjects designed to lessen this disparity—the right to an explanation of automated decisions—becomes feasible only from a "minimal informational" perspective. In other words, the informational asymmetry between users and operators of artificial intelligence systems introduces an additional layer of vulnerability for individuals in this context, as they lack access to the fundamental information necessary to articulate their grievances, make requests, and exercise their rights.

II. Scope and limitations of personal data protection laws

Several key issues have been raised by scholars regarding the general
inability of personal data protection laws to offer adequate protection to data subjects in situations analyzed\(^{22}\), particularly concerning the potential obfuscation between "algorithmic discrimination" and "algorithmic error" (stemming from "false positives" or "false negatives"). According to Schermer\(^{23}\), this scenario could unjustly place the burden of rights enforcement on the data subject. Furthermore, the safeguarding of trade secrets and industrial secrets emerges as a contentious issue.

Conversely, the LGPD (General Data Protection Law) introduces a robust principled framework concerning automated decisions. This framework includes principles like the Data Quality\(^{24}\) principle, which mandates that processed data be accurate, relevant, and current, thereby challenging non-representative profiling. It also embraces the Transparency\(^{25}\) principle, which guarantees data subjects the right to transparent, precise, and easily accessible information about data processing, offering a counter to "black box"\(^{26}\) tools. Notably, the Non-Discrimination\(^{27}\) principle forbids the use of data for unlawful or abusive discriminatory purposes, while the Accountability\(^{28}\) principle requires data controllers to demonstrate that they have implemented effective measures to ensure compliance.

Beyond these foundational principles, the LGPD provides specific provisions such as Articles 20 and 42, discussed in more detail in section 3.1 of this paper. These provisions are expected to be highly effective in resolving real-world conflicts and safeguarding the rights of data subjects as the Brazilian data protection landscape matures. The Brazilian National Data Protection Authority (ANPD), though still in its formative years, has actively engaged in sector regulation and fostered valuable synergy with initiatives aimed at regulating Artificial Intelligence in Brazil.\(^{29}\) Therefore, despite the

\(^{22}\) Schermer teaches that “given the rapid technological developments in the area of profiling, it is questionable whether the right to informational privacy and data protection law provide an adequate level of protection and are effective in balancing different interests when it comes to profiling.” Schermer, B. "Risks of Profiling and the Limits of Data Protection Law." In Discrimination and Privacy in the Information Society: Data Mining and Profiling in Large Databases, edited by Bart Custers et al. Heidelberg: Springer, 2013, 137-138.


\(^{24}\) Article 6º, V, LGPD.

\(^{25}\) Article 6º, VI, LGPD.


\(^{27}\) Article 6º, IX, LGPD.

\(^{28}\) Article 6º, V, LGPD. Referring to the duty of social agents to be accountable to controlling bodies.

\(^{29}\) ANPD, Coordenação-Geral de Tecnologia e Pesquisa. Nota Técnica n°
challenges identified by scholars in applying personal data protection laws within contexts of algorithmic discrimination, the outlook in the Brazilian scenario appears optimistic.

A. In between algorithmic discrimination and algorithmic error: data protection, and intellectual property

Profiling and algorithmic discrimination are integral concepts in discussions surrounding ethics, equity, and justice within artificial intelligence (AI). Profiling involves collecting and analyzing data on individuals to create profiles that accurately depict their characteristics, behaviors, and interests for a specific purpose. Algorithmic discrimination, on the other hand, occurs when AI systems unlawfully differentiate against individuals based on protected characteristics.

Although these concepts are interlinked, they are distinct, and their intersection often obscures the exercise of data subject rights. Schermer highlights that scenarios involving "false positives" or "false negatives" from algorithmic outputs are problematic because "it places the burden of proof on the side of the data subject: they must prove that they do or do not fit the profile." This burden is deemed disproportionate, especially in the context of vulnerability discussed in section 2.2 of this paper.

However, when considering the Brazilian General Data Protection Law (LGPD), Article 42, paragraph 2, offers a promising resolution to this challenge, should the conflict become judicialized. This provision allows judges in civil lawsuits the discretion to reverse the burden of proof in favor of the data subject under certain conditions, shifting the burden to the AI supplier or operator to prove the system's accuracy.

In non-judicialized conflicts, but within the administrative competence of the Brazilian National Data Protection Authority ("ANPD"), the provisions of Article 20 can provide excellent support. The Article's main provision grants the right to review automated decisions, and its paragraph 1 imposes the duty on the controller to provide "clear and adequate information regarding the criteria and procedures used for an automated decision, subject


to commercial and industrial secrecy." Anticipating the dispute between privacy and intellectual property rights, paragraph 2 gives the authority the prerogative to conduct an "audit to verify discriminatory aspects in automated processing of personal data" if this information is not provided for reasons of commercial and industrial secrecy.

Even in cases where the algorithm demonstrates a certain "accuracy" about the individual's profile, Mann and Matzner strengthen protection for data subjects by advocating for the principled application of antidiscrimination laws over personal data protection laws. The authors argue that "even if the algorithm in this case was not ‘wrong’ in a narrowly conceived epistemic understanding, applying the principles of antidiscrimination might mean refraining from using this information to make discriminatory decisions."

Schermer aligns with the idea that Anti-discrimination laws could be more efficient in this context, considering that fundamental institutes of personal data protection laws, such as the principle of minimization, may negatively impact the production of accurate AI tools. The author suggests the prevalence of anti-discrimination laws over personal data protection laws so that "data minimisation rules and prohibitions on the processing of sensitive data may be overruled if they undermine the accuracy of a profiling exercise, or if they deny us the possibility to detect discrimination in a profiling exercise." The problem posed by the author is that these systems are trained with large databases of personal data, while data protection laws vehemently foresee the processing of the smallest amount of personal data, and for AI systems, a shallow or unrepresentative base can be a major cause of biases.

Indeed, there seems to be, about the Principle of Minimization,

37 Which we infer, in the LGPD to the principle of Necessity, provided for in its Article 6, Section III.
interpretations that do not adequately connect with the foundations of personal data protection laws. The position of this author is that the Principle of Minimization does not support the use of insufficient—and possibly less expensive—databases to train algorithms of questionable quality. This is because its intention is not to treat a low numerical volume of data to the detriment of products and services secure to the privacy of the data subject. The principle's goal is that the processed data is adequate and relevant to the purpose for which it is intended, even on a large scale. In summary, if the purpose is to build a tool that is satisfactorily secure for the rights of the data subjects, the necessary data to design it can be processed without conflicting with this principle.

B. The exhaustive list of sensitive data and the contextual vulnerability

Primarily, for a brief perspective on the intersection between these themes, some concepts are especially important. Bruno Ricardo Bioni clarifies that sensitive data is categorized as such by the General Data Protection Law due to the possibility that its content offers special vulnerability to the respective data subject. Patrícia Peck Pinheiro explains that this category is usually related to "individual personality characteristics and personal choices" which the LGPD explicitly establishes as data regarding "racial or ethnic origin, religious belief, political opinion, union membership or organization of a religious, philosophical, or political nature, data concerning health or sexual life, genetic or biometric data when linked to a natural person."

Connecting to Bioni's concept, the Oxford English Dictionary defines "vulnerability" as "the fact of being weak and easily hurt physically or emotionally." These concepts relate to algorithmic discrimination, as

---

38 Referred to by the principle of “Necessity” provided for in Article 6, Section III of the LGPD.
39 Especially considering as one of the Foundations of the Brazilian General Data Protection Law “human rights, free development of personality, dignity and exercise of citizenship by natural persons”, according to art. 2nd, Section VII, of the LGPD.
40 It is also necessary to interpret it in synergy with the foundations set out in Article 2 of the LGPD, especially the provisions of Section V, from which we can infer that it is not the legislator’s intention that the provisions of the Law be incompatible with “economic and technological development and innovation”.
43 Article 5º, Section II, LGPD.
recruitment AIs are trained with databases from realities in which exclusionary structures prevail for certain social markers. A significant portion of these markers, subject to underrepresentation, can be expressed by sensitive data. In this sense, data subjects who possess them are more likely to suffer the damages caused by the inferences of a biased AI system, which will replicate these structures in an attempt to provide coherent outputs.

On the other hand, the exhaustive list proposed for the sensitive data category seems not to cover some factors that often transcend its subject to the same vulnerable position, such as female gender. This data category is often a dividing line for penalizing resumes of women, just as it happened in the unfortunate case of Amazon in 2015. This is a strong indicator that, as Florencia Luna proposes, vulnerability should be analyzed in layers - and not as labels - being highly contextual and experienced as "the result of the interaction of her particular circumstances and her own characteristics." Mann and Matzner argue that "intersectional theory has shown that safeguards against discrimination wrongly assume that all forms of discrimination function similarly or independently." Crenshaw argues that the combined effects of discrimination are particular forms of discrimination experienced by individuals with a specific combination of (protected) identities - and cannot be reduced to one of their "elements."

This is a particularly relevant premise for situations involving algorithmic discrimination, as the group of individuals (sharing a common characteristic) experiencing the effects of discrimination will vary due to both the criteria programmed into the AI and the database used for its training. Thus, which groups will be penalized by the algorithm's decision is not a factor that can be entirely labelled, but contextual, making it possible that the combination

---


of certain characteristics subjected to profiling techniques (layering vulnerabilities) may result in the construction of highly discriminated identities.

Furthermore, the vulnerability of the individual can occur without directly involving any sensitive data. That is, data not classified as sensitive by the privacy protection laws may equally place the individual in vulnerable positions. This can be observed in a widely publicized incident in the British media, exposing the fact that "an auto insurance, Ford Focus 2007 model, in the city of Leicester, was priced at 1,333 pounds for 'John Smith' and 2,252 pounds for 'Muhammed Ali'."\textsuperscript{49} That’s how some insurers set different values for applicants with the same profile based on discriminatory factors.

In the mentioned incident, it is noticeable that the AI system discriminately and unlawfully used user data, using their name to assume belonging to a certain group, which, according to its inferences, would have less favorable characteristics, justifying differential pricing. Such situations affect users in various positions in their social relations, whether as buyers, job applicants, applicants for certain financing, insurance, and so forth.

Both in the illustrated incident and the well-known Amazon recruitment episode, algorithmic discrimination occurred based on data that would not be categorized as sensitive under the Brazilian General Data Protection Law. The examples are sufficient to question whether an exhaustive list of sensitive data is suitable for its intended purpose or if AI regulation initiatives should progress to adequately encompass the dynamic and contextual nature of vulnerability.

III. INITIATIVES IN ARTIFICIAL INTELLIGENCE REGULATION

Regardless of the approach taken to mitigate the harmful effects caused by the vulnerability of workers as data subjects in recruitment processes, it is necessary to consider that the relevant AIs play significant operational roles in organizations. In this section, perspectives on regulatory advancements in the field will be addressed, such as the European I Act\textsuperscript{50}, the Brazilian Bill 2338/2023\textsuperscript{51}, and NYC Law 144\textsuperscript{52}, currently in effect in the state of New York.

\begin{footnotesize}


\textsuperscript{52} Zetoony, D. A., T. Boiangin, and F. Jerrold Goldberg. "NYC’s Local Law 144 and
York. These initiatives seek to reconcile innovation with the socially responsible use of AIs, especially in employment relationships.

As detailed in section 2 of this paper, personal data protection laws, by the very nature of their scope, do not cover all aspects that require regulation to prevent unlawful conduct in the processing of personal data by AI systems. Still, in connection with section 2, it is essential to note that the analyzed context highlights accumulations of layers of vulnerability on the individual, and corresponding protective mechanisms play a crucial role in achieving social justice.

In this sense, proposals for dedicated and specific regulation of the subject contribute significantly to advancing the discussion, both in Brazil and worldwide. The regulation of Artificial Intelligence can be a watershed moment in ensuring the protection of the rights of data subjects whose personal data is subjected to AI tools. By establishing clear guidelines for the appropriate use of AI, these regulations can consolidate the requirement for elements of algorithmic governance, contributing to fairness in selection algorithms.

In New York, the NYC Law 144, effective on July 5, 2023, specifically regulates the use of AI tools in recruitment. The law mandates that employers using algorithms in these contexts (recruitment, selection, and career movements) provide the necessary information for independent audits to make them public. Otherwise, there will be no permission for the use of AI in recruitment.\(^{53}\) Additionally, these companies must disclose potential scores used by the algorithm to classify candidates by race, ethnicity, and gender.\(^{54}\) When interpreted in conjunction with the discussion in section 1.3 of this paper, this requirement is understood as a protective measure corresponding to the layers of vulnerabilities that these social markers can impose on individuals.

In Europe, the process of enacting the IA ACT\(^ {55}\) is progressing, with its initiative and legal content inspiring various countries worldwide seeking to regulate aspects related to AI in their domestic legal frameworks. Regarding the intersection between artificial intelligence and the protection of personal data, this trend is expected to be even stronger, given that the European
A notable contribution of the AI Act is the classification of AI systems into risk levels. Depending on the category in which services and products are classified, agents would have a series of obligations related to privacy and transparency. These societal risk levels are divided into: a) low risk (exclusion criterion, applicable to tools not classified in other levels); b) limited risk (such as chatbots); c) high risk (including systems related to work and its management); and d) unacceptable risk (such as biometric systems for real-time surveillance of public spaces).

In the Brazilian Senate, since May 2023, the Bill 2338/2023 has been under consideration, with similar objectives, classifying AIs for "recruitment, screening, filtering, candidate assessment" as "High Risk" tools. To this category, the project establishes a series of governance obligations for AI agents, such as: i) Conducting Algorithmic Impact Assessment and sharing the assessment with the competent authority; ii) Developing technical documentation for AI systems before their market release or use in service provision, maintaining it during their use; and iii) Implementing measures to mitigate and prevent discriminatory biases.

The bill also introduces crucial provisions to counterbalance the layers of vulnerabilities that overlap individuals in the analyzed context. Article 12 proposes the explicit prohibition of the "use of artificial intelligence systems that may lead to direct, indirect, illegal, or abusive discrimination." Particularly concerning the prohibition of discrimination in its indirect form, it seems to align with the teachings of Mann and Matzner illustrated...
in section 3.1. That is, beyond analyses of algorithmic errors or discrimination, the automated processing of personal data should "refrain from using this information to make discriminatory decisions."\textsuperscript{62}

The proposal for Article 12, Section I, emphasizes its intention to establish itself as an additional counterbalancing measure against certain social markers that often represent layers of vulnerabilities for data subjects in recruitment processes. The provision explicitly grants data subjects the right to fair and equal treatment, prohibiting discrimination resulting from the "use of sensitive personal data or disproportionate impacts due to personal characteristics such as geographical origin, race, color, ethnicity, gender, sexual orientation, socioeconomic status, age, disability, religion, or political opinions."\textsuperscript{63} (emphasis added).

The exemplary nature of the prediction — rather than a restrictive one — expressly covering other personal characteristics subjected to disproportionate impact, is interpreted as favorable according to the line of reasoning defended in this paper. That is, it can be shaped to the multitude of characteristics susceptible to cause vulnerability to subjects and algorithmic discrimination. In the same perspective, the proposed Section II of the same article can serve as another strategy to counterbalance the effects of these layers. The text expressly prohibits the implementation and use of artificial intelligence systems that may lead to the "exacerbation of the vulnerability of people belonging to a specific group, even if apparently neutral criteria are used."\textsuperscript{64}

It is noted that its scope is broadened to encompass discriminatory decisions that i) are not strictly related to "sensitive data" as the inference is about a "specific group", which appears to be an advancement compared to the exhaustive list in Article 5, II of LGPD; and ii) are not the result of explicit and intentional programming in AI mechanisms, clarifying that the prohibition applies "even if apparently neutral criteria are used"\textsuperscript{65} which would be especially applicable to biases resulting from machine learning processes. Thus, considering the vulnerability of "specific groups" which certain social markers affect, and cumulatively, the condition of being an employee as well as a data subject, opaque, biased, and discriminatory AI tools certainly cause this "exacerbation of the vulnerability"\textsuperscript{66} the subject of


\textsuperscript{63} Article 12, Section I, Bill 2338/2023, freely translated.

\textsuperscript{64} Article 12, Section II, Bill 2338/2023, freely translated.

\textsuperscript{65} Article 12, Section II, Bill 2338/2023, freely translated.

\textsuperscript{66} Article 12, Section II, Bill 2338/2023, freely translated.
Ethical implementation of AI in job candidate recruitment

explicit prohibition.

In a systematic interpretation of the proposed text for Article 12, it can be observed that the legislator seeks to anticipate, through the provisions of Section I, that some social markers linked to the concept of sensitive data represent a significant portion of existing discriminations in our society. However, it also conceives that there are contextual and relational situations in which other "specific groups" are placed in a vulnerable position, justifying the provision of Section II. Systematic interpretation aligns with Luna's teachings that "there are situations that make individuals vulnerable and groups at greater risk than others."

As for the proposed text for Article 17 of Bill 2338/2023, it does not focus on counterbalancing vulnerabilities caused by specific social markers but rather on risks associated with the purposes of the AI tool itself. This approach seems to present an advance in mitigating the effects of the layers of vulnerability addressed in section 2 of this paper. The project, like the IA Act, categorizes AI tools intended for labor purposes as "high risk." In Section III of the same provision, the project specifically categorizes AIs with the purpose of "recruitment, screening, filtering, candidate assessment" into this category.

Another proposal of high relevance in the Bill is found in Article 27. Regarding "high-risk" tools, a category that includes recruitment AIs, its paragraph 1 proposes strict liability of the supplier or operator for damages caused, and its paragraph 2 provides for the presumed fault of the agent causing the damage, applying the "reverse the burden of proof in favor of the victim." This would present one more significant advance over the obstacle raised by Schermer in section 3 of this paper: the transfer to the data subject of the disproportionate burden of proving that they do not fit the profile defined by the AI.

In the Justification Section of the bill, it is clarified that it has a "dual objective", aiming at both: a) the "protection of the most vulnerable link in

67 The episode of algorithmic discrimination involving Amazon's recruitment AI in 2015 illustrates how specific groups, such as women, can be vulnerable to biases in these systems.
69 The proposed text for art. 17 of Bill 2338/2023 is highly relevant for the purposes of this article. It is cited: Art. 17. High-risk artificial intelligence systems are considered to be those used for the following purposes: “(...) III – recruitment, screening, filtering, evaluation of candidates, making decisions about promotions or termination of relationships employment contracts, division of tasks and control and evaluation of the performance and behavior of people affected by such artificial intelligence applications in the areas of employment, worker management and access to self-employment”, freely translated.
70 Article 17, Section III, Bill 2338/2023, freely translated.
71 Article 27, paragraph 2, Bill 2338/2023, freely translated.
question"72 which is "the natural person who is already daily impacted by artificial intelligence systems"73 and b) the provision of "governance tools and an institutional arrangement for monitoring and supervision"74, which would create "conditions of predictability regarding its interpretation and, ultimately, legal certainty for innovation and technological development"75. It is noted that the justification for its proposal is illustrative in summarizing the objectives guiding Brazil and other countries around the world to regulate artificial intelligence, mitigating the risks associated with its use for socially responsible purposes.

In a combined interpretation of the illustrated initiatives, it can be concluded that there is a global tendency to perceive recruitment AIs as high-risk systems, and furthermore, to classify risk categories for other types of AIs, seeking to establish proportional governance obligations for the agents. In general, the worker, as a personal data subject, tends to be safeguarded with strategies coherent with their position of vulnerability, in addition to protections related to specific social markers. Global initiatives aim to establish AI suppliers and operators as primary responsible parties for ensuring their ethical use and development, associated with accountability duties. This is an optimistic scenario and, in synergy with personal data protection laws, tends to mitigate the obstacles raised by doctrine regarding the realization of data subject rights.

CONCLUSION

As technological advancement and artificial intelligence (AI) continue to reshape social relations, concerns grow about how to reconcile technological development with ethical and responsible parameters in its use. Brazil, along with many other countries worldwide, has committed to legislative initiatives that address the risks associated with the contextual purposes of these tools and the layers of vulnerability overlapping with data subjects.

For the purposes of this research, it is concluded that there is a trend towards initiatives related to the subject, considering that: a) Employment relationships have special considerations regarding the use of artificial intelligence since this is inherently an asymmetrical relationship, which should not be exacerbated to the detriment of the data subject; b) The majority of users whose data are processed by these tools are in a relationship of informational asymmetry with AI agents, fostering a duty of transparency and

72 Justification Section, Bill 2338/2023, freely translated.
73 Justification Section, Bill 2338/2023, freely translated.
74 Justification Section, Bill 2338/2023, freely translated.
75 Justification Section, Bill 2338/2023, freely translated.
information on the part of the agents; c) Workers, as data subjects, may be subject to additional layers of vulnerability arising from certain social markers, which also deserve corresponding mitigation strategies.

In this context, the conflict between automated decisions and the right to transparency and non-discrimination highlights the need for effective regulations to safeguard the legal rights of data subjects beyond the scope of data protection laws. National and international legislative initiatives in the field of artificial intelligence underscore the global trend towards regulation aimed at ultimately reconciling technological development with the ethical, transparent, and responsible use of AI in social relations.

REFERENCES


Parentoni, L. "What Should We Reasonably Expect From Artificial Intelligence?" ResearchGate, 2022. Available at: https://www.researchgate.net/publication/361988480_What_should_we_reasonably_expect_from_artificial_intelligence.


* * *

Jessica Fernandes Rocha
Email: jrocha@stoccheforbes.com.br
ORCID iD: https://orcid.org/0009-0003-4985-9086

doi: 10.59224/bjlti.v2i1.120-139