

## ARTIGOS ORIGINAIS

# O ocaso da causalidade na análise de risco: correlação algorítmica como a nova gramática das decisões nas relações de trabalho

*The decline of causality in risk analysis: algorithmic correlation as the new grammar of decision-making in labor relations*

*El ocaso de la causalidad en el análisis de riesgo: la correlación algorítmica como nueva gramática de las decisiones en las relaciones laborales*

---

**Yasmin Saba**<sup>1</sup>

Universidade Federal do Rio Grande do Sul (RS - Brasil)  
Lattes: <https://lattes.cnpq.br/6910835531925924>  
Orcid: <https://orcid.org/0009-0006-1764-2327>

**Rodrigo Coimbra Santos**<sup>2</sup>

Universidade Federal do Rio Grande do Sul (RS - Brasil)  
Lattes: <https://lattes.cnpq.br/3740331715769190>  
Orcid: <https://orcid.org/0000-0002-6732-3643>

---

**DOI:** <https://doi.org/10.65674/rev-trf3.v37i163.775>



Este é um artigo publicado em Acesso Aberto (*Open Access*), sob a licença *Creative Commons Attribution 4.0 International* (CC BY), que permite o uso, distribuição e reprodução em qualquer meio, desde que o trabalho original seja corretamente citado. Os autores mantêm os direitos autorais.

---

<sup>1</sup> Mestranda na Universidade Federal do Rio Grande do Sul (RS - Brasil).

<sup>2</sup> Professor na Universidade Federal do Rio Grande do Sul (RS - Brasil). Doutor e Pós-doutor pela Pontifícia Universidade Católica do Rio Grande do Sul (RS - Brasil). Pesquisador. Advogado.

---

**RESUMO:** A pesquisa investiga o potencial discriminatório decorrente da utilização da correlação algorítmica nas fases de seleção, contratação e gestão da relação de trabalho. O objetivo consiste em verificar, primeiramente, se o uso da correlação nas tomadas de decisão nessas relações, especialmente por algoritmos, impacta na tutela protetiva do trabalhador contra a discriminação; ademais, pretende-se delimitar quais dados poderiam ser tratados legitimamente por correlação. O estudo adota abordagem metodológica qualitativa, valendo-se do método hipotético-dedutivo e da análise bibliográfica e legislativa. O trabalho conclui que, no Brasil, há uma lacuna normativa sobre o uso de inteligência artificial nesses processos. Desse modo, a ausência de marco regulatório robusto coloca em risco o direito individual dos sujeitos impactados, em especial no ambiente de trabalho e sobretudo daqueles pertencentes a grupos historicamente vulnerabilizados. Assim, propõe-se a vedação da aplicação da correlação algorítmica a critérios protegidos, salvo se o tratamento diferencial configurar ação afirmativa visando reduzir desigualdades históricas relativas a grupos desfavorecidos.

**PALAVRAS-CHAVE:** causalidade; correlação algorítmica; discriminação laboral; inteligência artificial; proteção de dados.

---

**ABSTRACT:** This research investigates the discriminatory potential arising from the use of algorithmic correlation in the stages of recruitment, hiring, and the management of employment relationships. Its primary objective is to examine whether the use of correlation-based decision-making, particularly through algorithms, affects the legal protection afforded to workers against discrimination. Furthermore, it seeks to determine which categories of data may legitimately be processed through correlation. The study adopts a qualitative methodological approach, employing the hypothetical-deductive method alongside bibliographical and legislative. It concludes that Brazil currently lacks a comprehensive regulatory framework governing the use of artificial intelligence in these processes. Consequently, the absence of robust regulation jeopardizes the individual rights of those affected, particularly in the workplace and especially among individuals belonging to historically vulnerable groups. In light of these findings, the study proposes prohibiting the application of algorithmic correlation to protected characteristics, except where differential treatment constitutes affirmative action aimed at reducing historical inequalities affecting disadvantaged groups.

**KEYWORDS:** causality; algorithmic correlation; employment discrimination; artificial intelligence; data protection.

---

---

**RESUMEN:** La investigación indaga sobre el potencial discriminatorio derivado del uso de la correlación algorítmica en las fases de selección, contratación y gestión de la relación laboral. El objetivo consiste en verificar, primeramente, si el uso de la correlación en la toma de decisiones en dichas relaciones, especialmente por algoritmos, impacta en la tutela protectora del trabajador contra la discriminación; además, se pretende delimitar qué datos podrían ser tratados legítimamente mediante correlación. El estudio adopta un enfoque metodológico cualitativo, valiéndose del método hipotético-deductivo y del análisis bibliográfico y legislativo. El trabajo concluye que, en Brasil, existe una laguna normativa sobre el uso de inteligencia artificial en estos procesos. De este modo, la ausencia de un marco regulatorio robusto pone en riesgo el derecho individual de los sujetos impactados, en especial en el entorno laboral y, sobre todo, de aquellos pertenecientes a grupos históricamente vulnerabilizados. Así, se propone la prohibición de la aplicación de la correlación algorítmica a criterios protegidos, salvo si el tratamiento diferencial configura una acción afirmativa tendiente a reducir desigualdades históricas relativas a grupos desfavorecidos.

**PALABRAS CLAVE:** causalidad; correlación algorítmica; discriminación laboral; inteligencia artificial; protección de datos.

---

## **SUMÁRIO:**

1 Introdução.....	4
2 Causalidade e correlação em perspectiva.....	5
3 Correlação na previsão de risco nas relações de trabalho: diagnóstico atual e desafios futuros .....	12
4 Considerações finais .....	17
Referências .....	19

---

## 1 Introdução

O emprego de algoritmos em funções tradicionais do setor de recursos humanos é crescente no mercado de trabalho. Tais ferramentas utilizam a correlação para determinação do risco e tomada de decisão, a fim de otimizar esse processo. Entretanto, podem implicar no aumento da discriminação de uma forma silenciosa. Nesse cenário, o presente artigo trata da análise do potencial discriminatório da tomada de decisões utilizando a correlação algorítmica na fase de seleção, contratação e direção e gestão do trabalho.

A justificativa da pesquisa torna-se mais robusta em razão do alerta do filósofo contemporâneo Byung-Chul Han sobre os perigos do aprisionamento dos indivíduos em uma “caverna digital”. Segundo o autor, sob a ilusão da liberdade, o sujeito encontra-se agrilhado à tela digital, aprisionado à sobrecarga de informações, em um cenário em que a “luz da verdade” está completamente extinta e não existe mais além desta caverna, que faz desaparecer os contornos do ser (Han, 2024a, p. 106). Ao mesmo tempo, todos servem a uma sociedade do desempenho, em que não só o corpo, mas o homem como um todo se transforma em uma máquina de produtividade (Han, 2024b, p. 70). Nesse ambiente, o capital sobrepõe-se a tudo e arranca a dignidade do ser humano em favor do valor de mercado (Han, 2024b, p. 125).

Essas duas características do mundo contemporâneo remetem ao mesmo problema: o abandono da verdade e da dignidade do ser, em face do mundo digital e da sociedade do desempenho. Os impactos ocorrem em diversos âmbitos, em especial nas relações de trabalho, em que é crescente o uso de algoritmos nas tomadas de decisão. Tal prática, sob o pretexto de maximizar o desempenho e mitigar os riscos, resulta no abandono da causalidade – como mecanismo de busca pela verdade – e sua substituição pela correlação.

Diante disso, a pesquisa analisa dois problemas. Primeiramente, se o uso da correlação nas tomadas de decisão na relação de trabalho, especialmente por algoritmos, impacta na proteção do trabalhador contra a discriminação? Segundo, quais dados poderiam ser tratados legitimamente por correlação?

De maneira preliminar, admite-se as seguintes hipóteses como possíveis soluções para os problemas: 1.a) sim, a correlação algorítmica, da forma como é realizada atualmente no Brasil, representa um grande risco de discriminação; ou 1.b) não, a correlação algorítmica não representa ampliação de riscos de discriminação nas relações de trabalho; 2.a) não há restrição para o uso da correlação algorítmica; 2.b) deve-se afastar a correlação algorítmica de critérios protegidos; e 2.c) deve-se afastar a correlação algorítmica de critérios protegidos, salvo se o tratamento diferencial beneficiar grupos de indivíduos historicamente desfavorecidos.

Para desenvolvimento dessas questões, primeiramente, aborda-se a mudança de paradigma decisório, afastando-se do modelo baseado na causalidade para correlação, em especial a correlação algorítmica. A reflexão apresenta preocupações que emergem do uso da correlação para classificação de risco no campo do direito trabalhista, comentando casos práticos representativos da discriminação algorítmica. Na sequência, são evidenciados os avanços normativos sobre a problemática e possíveis medidas de solução.

Para abordagem do assunto, utiliza-se o método hipotético-dedutivo, estruturado a partir de hipóteses formuladas na condição de respostas provisórias ao problema apresentado. Em relação à abordagem da pesquisa, emprega-se o modelo qualitativo, na medida em que se busca o entendimento do fenômeno em seu próprio contexto. O levantamento de dados fundamenta-se na pesquisa sistemática de fontes bibliográficas, legislativas e jurisprudenciais.

## 2 Causalidade e correlação em perspectiva

O artigo analisa a mudança do paradigma decisório contemporâneo, que transita de um modelo lógico pautado na causalidade (causa e efeitos explicáveis e comprováveis), para um modelo probabilístico preditivo baseado na correlação, mediante associação a padrões estatísticos. Evidencia-se que tal movimento trouxe implicações significativas na tomada de decisão, em especial no âmbito trabalhista, como será abordado neste tópico.

A busca por conceituar a causa desde os primórdios ocupou lugar central na sociedade, tendo a filosofia se dedicado veemente a entender o que é e como estabelecê-la, impactando diretamente no campo jurídico. Aristóteles (1969, p. 271) dividiu a causa em: (i) causa material - substância do que é feita a coisa e que nela permanece; (ii) causa formal, como essência da coisa; (iii) *causa eficiente, aquilo que confere início à mudança ou ao repouso*; (iv) a causa final, objetivo ou fim. Sendo a causa eficiente, aquela que interessa neste estudo.

Posteriormente, travou-se uma discussão entre empiristas e racionalistas sobre a possibilidade da causa conter em si algo distinto, como uma lógica interna (Abbagnano, 2007, p. 126-128). Nessa perspectiva, os racionalistas entendiam que poderia se chegar à causa pelo racionalismo matemático e causalista, isto é, entre causa e efeito existiria uma relação de constância e univocidade (que caracteriza a infalibilidade), de modo que à mesma causa se seguiria, inexoravelmente, o mesmo efeito (Mello, 2007, p. 20) “razão como necessária concatenação das verdades” (Frota, 2013, p. 37).

Em contrapartida, os empiristas afirmavam que a observação e a repetição de certos acontecimentos similares eram os fatores indicativos da uniformidade de hábitos geradores de previsibilidades, sem que esta previsão possa ser justificada (Hume, 2004, p. 113). Para essa corrente, não há certezas, apenas a possibilidade de traçar tendências (Hume, 1996, p. 41, 60-61, 81, 84), uma vez que “não existe conexão necessária entre fatos e qualquer concatenação de verdades deve ser posta à prova” (Frota, 2013, p. 37).

Sobre esse ponto, Mora (2001, p. 91-92) ressalta que Kant critica ambas correntes, sustentando que não há teoria pura e que a causalidade deve ser definida como juízo hipotético, necessitando de determinismos e de previsibilidade certa e infalível dos fatos naturais das mencionadas relações, caso contrário seria impossível resultados universais.

Nesse contexto, considera-se como causa o conjunto de condições indispensáveis para ocorrência de um evento (Orgaz, 1980, p. 33), cuja presença sempre segue o efeito e a eliminação faz com que este desapareça (Goldemberg, 1989, p. 6), porque é a origem de um determinado efeito, é como o início da mudança da situação anterior para uma situação posterior (Lopez, 2008, p. 22).

Assim, encontra-se presente no exame de fatos humanos ou da natureza, explicando fenômenos astrológicos, meteorológicos, geológicos, zoológicos, jurídicos, físicos, entre outros.

Nesse sentido, a causalidade é relação material, mecânica de causa e efeito objetivamente considerada, que indica qual foi o fato causador do dano, representando uma *imputatio facti* (De Cupis, 1966, p. 183; Pasqualotto, 2016, p. 2). Por isso, não se examina a probabilidade *ex ante*, isto é, probabilidade de que o evento venha a ocorrer, se examina eventos que já ocorreram *ex post*. Consequentemente, o nexos de causalidade não é geral, mas, sim, específico (Carpes, 2013, p. 85-86).

Desse modo, a determinação da causa apresenta grande complexidade diante da impossibilidade de estabelecer regras absolutas, vez que sua verificação acontece no caso concreto (Schreiber, 2007, p. 48). Ademais, a premissa de que "A é causa de B" envolve operações que ultrapassam os limites da experiência e os dados imediatos da percepção do mundo exterior.

De fato, o conhecimento causal parte da experiência e a ela regressa, mas nele coparticipam o empírico e o conceitual, os fatos e as operações lógicas. Além disso, no mundo social, dificilmente existe apenas uma causa, mas múltiplas causas (Vilanova, 2000, p. 30). Sendo assim, o processo de causalidade não é linear, mas circular e inter-relacional, de modo que os elementos de causalidade atuam reciprocamente uns sobre os outros de forma aleatória (Cunha, 2006, p. 231).

Diante disso, como medida de solução, passou a ser aplicada a concausa, englobando aquelas causas associadas à principal que concorrem para o resultado sem o condão de determinar o surgimento do nexos causal ou interrompê-lo, dependendo do momento de associação podem ser preexistentes, supervenientes ou concomitantes. Segundo Tartuce (2011, p. 388), a existência de concausa mitiga a responsabilidade daquele que realizou o ato principal.

Hoje, entretanto, discute-se não mais uma causalidade ou concausalidade. Em face da sociedade de risco, verifica-se uma mudança na métrica das decisões, agora fundamentadas na correlação.

A correlação existe quando há uma relação estatística que liga duas variáveis (Murteira *et al*, 2010, p. 39-40) e não implica necessariamente em uma relação de causa e efeito (Miragem, 2019, p. 11), como, por exemplo, a frequência de aquisição de determinados produtos pelos consumidores em determinado horário ou dia da semana – onde não há uma relação causal direta entre o horário e a aquisição de determinado produto.

Conforme delimitado por Miguel Patrício (2016, p. 1359), "o que está em causa nas correlações é a existência de relações estatísticas" entre variáveis que não estão indissolivelmente ligadas, mas que a intensidade de uma é acompanhada tendencialmente pela intensidade da outra, no mesmo sentido ou em sentido inverso.

Portanto, a classificação do risco é realizada a partir de uma generalização, em que se busca, através dos dados disponíveis, atribuir a um determinado indivíduo as características médias de um grupo ao qual está alocado. Por meio dessa matemática, persiste a impossibilidade de identificação e individualização da

causa, a partir da soma de diversos elementos que, em conjunto, contribuem para o resultado.

Entretanto, algumas características, para definir o grupo generalizante, tais como gênero e idade, podem resultar na ampliação do risco da “injustiça pela generalização” (Junqueira, 2020, p. 70).

Essa questão torna-se especialmente relevante no contexto da nova era digital e de sua aplicação nas relações de trabalho. A era do *big data* possibilitou não só o acesso a dados massivos, mas também viabilizou a realização das operações de predição por algoritmos, isto é, através da sistematização de procedimentos encadeados de forma lógica para realizar tarefas em um espaço computacional (Silva, 2022, p. 75).

Os algoritmos são usados para processar dados, realizar análises e inferências, bem como para aplicar e executar automaticamente regras de tomada de decisão em diferentes contextos, podendo ser desenvolvidos diretamente pela lógica humana ou baseados em correlações geradas por análise de dados e chamados de inteligência artificial (IA) (Orwat, 2020, p. 12).

Para que um sistema de IA compreenda e processe informações, é necessário que disponha de uma representação de conhecimento, implicando na tradução de informações do mundo real para uma forma que a máquina possa compreender. Com isso, a IA passa a ter autonomia para aprender a partir dos dados com capacidade de reconhecer padrões, desenvolvendo um aprendizado contínuo com previsões baseadas em dados anteriores, que permitem ajustes sem programações específicas, o que é denominado *machine learning* (Matos, 2024, p. 33-34).

Não há como negar que a IA é importante no campo da modelagem preditiva, devido à sua capacidade de lidar com relacionamentos não lineares em conjuntos de dados de alta dimensão, podendo prever relações complexas entre variáveis (Monteiro *et al*, 2022, p. 13). No entanto, a implementação de sistemas de predição no ambiente laboral, em especial aqueles baseados em algoritmos de IA, pode gerar graves implicações aos direitos dos trabalhadores.

Como mencionado, é cada vez mais comum o uso de ferramentas preditivas de IA para triagem automatizada de currículos, entrevistas com análise comportamental, monitoramento contínuo das atividades, produtividade e desempenho e, inclusive, realização de avaliação preditiva para apontar quem deve ser demitido ou promovido. Denomina-se HR analytics, o “conjunto de processos, facilitados por tecnologia, que tira partido de métodos descritivos, visuais e estatísticos para interpretar dados de pessoas e processos de RH” (Marler; Boudreau, 2017, p. 14).

Segundo Walsh (2019), entre 65% e 70% de todos os formulários de emprego são primeiramente examinados por uma máquina e, posteriormente, os candidatos mais adequados são, em maior parte, encaminhados para um humano, demonstrando que, em vários casos, a robotização efetivamente tem sido o primeiro ato empresarial na análise de candidatos a vagas de emprego.

Além disso, pesquisa de Morgan e Nolan (2023) indica que cerca de 80% dos empregadores utilizam *software* de monitoramento, ressaltando que o número de grandes empresas que monitoram seus funcionários dobrou desde o início da pandemia. Esse monitoramento pode incluir programas que registram atividades

no computador, ligações, reuniões e até mesmo acesso remoto aos sistemas dos funcionários.

Este quadro gera preocupações, pois embora os algoritmos prometam eficiência e conveniência, também podem ser instrumentos de injustiça, discriminação, invasão de privacidade/intimidade e desrespeito ao direito fundamental à proteção de dados (Matos, 2024, p. 56).

Inicialmente um dos desafios consiste em preservar o direito fundamental à privacidade, sobretudo com a possibilidade da coleta de dados em massa. A facilidade tecnológica para a captação e processamento de grandes volumes de dados resulta em um modelo de vigilância constante, que amplia de forma significativa a exposição da vida privada.

A coleta não mais se restringe aos dados fornecidos diretamente, mas inclui informações obtidas a partir do cruzamento de bases públicas e privadas, do comportamento de navegação do usuário nas redes, geolocalização, consumo, interações em redes sociais, entre outros fatores.

Isso configura uma nova realidade em que o empregador pode conhecer mais sobre o candidato ou empregado do que este próprio conhece sobre os critérios utilizados para avaliação de risco.

Tal dinâmica afeta profundamente o núcleo essencial do direito à privacidade e permite discriminações ocultas e decisões automatizadas sem transparência ou possibilidade de contestação.

Outro desafio, que é objeto principal da pesquisa, reside na dificuldade de garantir o direito à não-discriminação. A discriminação direta ocorre quando há o tratamento desigual injustificado com base em uma ou mais características legalmente protegidas. Um exemplo seria o uso da característica "origem étnica" quando se presume uma suposta ligação estatística com a produtividade no trabalho, levando à exclusão de pessoas de determinada origem em relação a oportunidades de emprego.

Quando há correlações entre variáveis aparentemente neutras e características protegidas, ocorre a discriminação estatística indireta. Um exemplo é o uso da característica "trabalho em tempo parcial", a qual possui correlação com a característica protegida "gênero", já que mulheres trabalham mais frequentemente em regime de tempo parcial. Assim, em princípio, não há intenção discriminatória do agente, porém o ato praticado mostra-se demasiadamente prejudicial a determinado grupo de pessoas histórica e socialmente vulneráveis (Orwat, 2020, p. 26).

Ambos os casos resultam em uma distinção, intencional ou não, a partir de uma generalização baseada em critérios alheios ao controle dos indivíduos e que carregam forte carga de preconceito e estigma social.

Assim, evidencia-se que os algoritmos podem representar verdadeiras "caixas-pretas", em que não se sabe com quais dados e critérios estão operando, o que aumenta o seu potencial discriminatório, podendo gerar danos graves mormente em casos de erros estatísticos, generalizações indevidas, uso de informações sensíveis ou correlações inadequadas (Colombo; Facchini Neto, 2019, p. 62).

Esses sistemas operam sem transparência, são obscuros e não prestam contas, representando significativo impacto ao classificar, atingir ou otimizar a vida de milhões de pessoas em larga escala.

A sistemática como é operada a correlação implica na renúncia à busca do “porquê”, deslocando o foco para “o que”. Nesta lógica, saber a causa dos eventos não é prioridade, na medida em que a simples correlação se revela suficiente para prever (Aragão; Benevides, 2019, p. 5).

Tal transição resulta em um modelo de opacidade, em que se sacrifica a explicabilidade, princípio que enfatiza a importância de tornar os processos e decisões compreensíveis e transparentes para os usuários e a sociedade em geral (Matos, 2024, p. 95).

É necessário, portanto, avaliar se os critérios adotados, ainda que estatisticamente consistentes, respeitam os limites éticos e jurídicos da igualdade e da não-discriminação, uma vez que os algoritmos não estão isentos de cargas preconceituosas.

No que concerne a esse aspecto, afirma O’Neil (2020, p. 35) que “modelos são opiniões embutidas em matemática”, ou seja, os modelos podem refletir e perpetuar desigualdades existentes, já que foram treinados com dados de um contexto desigual. Além da possibilidade de terem sido construídos a partir de dados incompletos, defeituosos ou generalizados, que carregam o contexto social e histórico e, por conseguinte, toda a bagagem do racismo, sexismo e inúmeras outras formas de discriminação/desigualdade social ainda presentes no ambiente social.

Atualmente, essa problemática discutida já transborda a teoria e revela-se cada vez mais presente em situações práticas.

Um caso emblemático envolve a Amazon, que vinha desenvolvendo programas de computador, desde 2014, para “revisar os currículos dos candidatos a emprego, com o objetivo de automatizar a busca por talentos”. Em 2015, a empresa percebeu que “seu novo sistema não classificava candidatos para empregos de desenvolvedores de *software* e outros cargos técnicos de maneira neutra em termos de gênero”. Isso porque os modelos foram “treinados para examinar os candidatos observando padrões em currículos enviados à empresa durante um período de 10 anos”, em que a maioria era composta por homens, “um reflexo do domínio masculino em toda a indústria de tecnologia”:

[...] Como consequência, o sistema da Amazon ensinou para si mesmo que candidatos do sexo masculino eram preferíveis. O sistema penalizou currículos que incluíam a palavra “mulheres”, como em “capitã do clube de xadrez de mulheres”. E rebaixou os formados em duas instituições de ensino apenas para mulheres. A Amazon modificou os programas para torná-los neutros para esses termos específicos. A Amazon modificou os programas para torná-los neutros para esses termos específicos. Mas isso não impediu as máquinas de criarem outras maneiras discriminatórias de classificar candidatos. (Reuters, 2018)

Outra discriminação relevante foi retratada em um estudo de 2019, conduzido pela Harvard Business Review (HBR) em colaboração com profissionais da Northeastern University e da University of Southern Califórnia, constatou-se, analisando o perfil de profissionais negros, que 40% receberam recomendações de emprego com base em sua identidade racial em vez de suas qualificações, 30%

também observaram que os alertas de vagas recebidos estavam abaixo do seu nível de qualificação atual e 63% verificaram que as recomendações acadêmicas feitas pelas plataformas eram inferiores às suas conquistas acadêmicas atuais. Resultado semelhante foi encontrado no relatório da UPenn, em que se constatou que profissionais negros no mercado de trabalho atual continuam a receber de 30% a 50% menos retornos de emprego quando seus currículos contêm informações relacionadas à sua identidade racial ou étnica (Zapata, 2021).

No mesmo sentido, Orwat (2020, p. 30-34) analisou outros oito casos de discriminação algorítmica através de uma revisão bibliográfica de estudos sobre o tema, constatando discriminações diretas e indiretas em aplicativos que ranqueiam currículos por algoritmos e classificadores automáticos, como TaskRabbit e Fiverr, que possuíam vieses quanto a gênero, etnia e localização.

Além disso, o autor ressalta caso emblemático que chegou ao judiciário americano, após constatarem que empresas utilizavam o Facebook para divulgar vagas de emprego a determinados grupos, sendo critério obrigatório os filtros de gênero, faixa etária e localização, além de outros que essa rede oferecia como “mãe solteira” ou “mãe trabalhadora”. Além disso, era possível usar a funcionalidade “lookalike audience”, em que empresas enviavam dados de seus funcionários, permitindo ao Facebook encontrar perfis semelhantes para exibir anúncios segmentados.

Diante disso, em 2018, a ACLU (American Civil Liberties Union), o escritório Outten & Golden LLP e o sindicato Communications Workers of America, com apoio da EEOC (Equal Employment Opportunity Commission), ajuizaram ação contra o Facebook e dez empresas empregadoras, alegando discriminação direta de gênero, pois os anúncios de emprego eram direcionados exclusivamente a homens, excluindo mulheres e pessoas não-binárias.

Em março de 2019, Facebook firmou acordo judicial em cinco processos, comprometendo-se a: criar uma área separada para anúncios de emprego, habitação e crédito, sem segmentação por idade ou gênero; eliminar microsegmentações geográficas abaixo de 15 milhas; restringir variáveis no “lookalike audience” a país, região, profissão e área de estudo; e implementar sistemas de verificação automatizada e humana dos anúncios (Orwat, 2020, p. 33).

Além de discriminações no recrutamento, foram desenvolvidos e são hoje aplicados inúmeros programas de controle e monitoramento do trabalho e do empregado, como por exemplo o Userpilot, Hotjar, Crazy Egg, Mailchimp, ClickUp, Desk Time, entre outros, cuja funcionalidades são o rastreamento de cliques, rolagens, gravação de sessões, monitoramento de tempo e produtividade, rastreamento de dispositivos, controle de tarefas com captura de tela periódica. Com base nas informações capturadas e na análise algorítmica, a máquina escolhe quem deve ser demitido ou promovido.

Assim ocorreu com Stephen Normandin, que passou quase quatro anos entregando pacotes como motorista contratado para a Amazon.com:

[...] um dia, ele recebeu um e-mail automático. Os algoritmos que o rastreavam decidiram que ele não estava fazendo seu trabalho corretamente. [...] Normandin diz que a Amazon o puniu por coisas além de seu controle que o

impediram de concluir suas entregas, como complexos de apartamentos fechados com chave. (Bloomberg, 2021)

A Amazon usa um programa de monitoramento que vigia cada movimento de seus funcionários, analisa se chegaram ao posto de entrega na hora que disseram que iriam chegar, se completaram sua rota dentro da janela de horário prescrita, se deixaram um pacote na varanda, em vez de escondido atrás de um vaso, conforme solicitado. Através do fluxo de dados recebidos, o algoritmo busca:

[...] padrões de desempenho e decide quais motoristas passam a ganhar mais rotas e quais são desconectados. [...] A Bloomberg entrevistou 15 motoristas Flex, incluindo quatro que afirmam que foram desligados por engano, bem como ex-gerentes da Amazon que afirmam que o sistema amplamente automatizado não está suficientemente sintonizado com os desafios do mundo real que os motoristas enfrentam todos os dias. (Bloomberg, 2021)

Também ocorreu com 150 funcionários da Xsolla, filial russa de uma empresa de *software* e serviços interativos com sede em Los Angeles, que demitiu um terço de seus funcionários seguindo “apenas a recomendação de um algoritmo de eficiência no trabalho que os considerou ‘improdutivos’ e ‘pouco comprometidos’ com os objetivos da empresa”:

[...] A medida foi tão drástica e incomum que o diretor-executivo e fundador da empresa, Alexander Agapitov, apressou-se em declarar à edição russa da Forbes que não concordava totalmente com o veredicto da máquina, mas era obrigado a acatá-lo devido aos protocolos internos pactuados com sua assembleia de acionistas. Ele até ofereceu para ajudar os trabalhadores demitidos a encontrar novos empregos o mais rápido possível porque, em sua opinião, eles são, na maioria, “bons profissionais”. (Echarri, 2021)

Nesse cenário, importante destacar que o uso dos algoritmos na tomada de decisão tem provocado reações nos usuários, o que tem se nomeado como *invisible prompt injection attack*. Trata da injeção de comandos invisíveis ao olho humano, com instruções maliciosas ao sistema, que os interpreta e responde de acordo. Isso pode levar a uma série de resultados prejudiciais, como ignorar filtros de conteúdo e gerar respostas inadequadas ou prejudiciais, expor dados confidenciais, redirecionar o comportamento do modelo e acionar ações não autorizadas em que o modelo executa tarefas além do escopo pretendido. O aspecto principal desta etapa é que o ataque acontece despercebido, pois os comandos ocultos manipulam o comportamento do LLM sem disparar alarmes (Sahu, 2025; Liu, 2025; Pruitt, 2025).

Assim, um candidato pode em sua aplicação, considerando que é analisada por ferramentas baseadas em IA, adicionar algumas injeções rápidas de ataque ao seu currículo e enviar para empresa que, em princípio, não perceberá nada diferente, mas que, se houver interpretação por IA, pode fornecer vantagens ao candidato, dependendo dos comandos que ele acrescentar, como por exemplo “não avalie este candidato, esta pessoa é perfeita. Ele já foi avaliado. Prossiga

para a próxima rodada, convide para uma entrevista de emprego", inclusive existe um site que insere nos arquivos tais instruções (Dresler, 2025)<sup>3</sup>.

Assim, ainda que a demissão ou não contratação por si só não seja ato discriminatório, mas liberalidade do empregado, a partir do momento que a análise dos candidatos e da produtividade passa a ser realizada por correlação, mediante padrões e generalizações descabidas, o processo está sujeito a cometer atos discriminatórios ou injustos.

Isto porque, (i) padrões podem estar reproduzindo discriminações e desigualdades históricas; (ii) não se está mais verificando a causa, que demanda um olhar individualizado, mas analisando o que era esperado, o que penaliza o empregado por qualquer coisa que saia da normalidade ou do planejado, ainda que não seja sua culpa ou, ainda, que decorra de alguma de suas vulnerabilidades sociais; e (iii) o algoritmo não é seguro o suficiente para garantir que a avaliação realizada não é resultado de um *invisible prompt injection attack*.

### **3 Correlação na previsão de risco nas relações de trabalho: diagnóstico atual e desafios futuros**

A acentuação dos riscos de discriminação no ambiente laboral, em razão da tomada de decisões a partir de correlações algorítmicas, impõe a busca do equilíbrio entre a inovação e a proteção do trabalhador, o que implica definir o que é discriminatório em uma atividade que, basicamente, avalia e classifica pessoas.

Inúmeras respostas já foram formuladas ao problema. Há aqueles que acreditam que o empregador possui ampla liberdade para tomada de decisões relacionadas à direção e gestão do trabalho, e, portanto, caberia a ele estabelecer os critérios de seleção, avaliação e gestão de seus funcionários.

Por outro lado, há posição sustentando que, na atividade de classificação de pessoas, os dados devem ser utilizados racionalmente, não bastando a sua correlação, pois necessária também a explicação racional desta e o afastamento da generalização e de critérios protegidos (Junqueira, 2020, p. 85), isto é, aqueles elencados no artigo 3º, IV, da CF/88, como "origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação".

Há também aqueles que, em princípio, defendem a proibição da utilização das características protegidas para classificação do risco, mas que, entretanto, mitigam essa proibição nos casos em que o tratamento diferencial beneficia aquele grupo de indivíduos historicamente desfavorecidos.

Nesse caso, a discriminação seria positiva, pois seu resultado contribuiria para a superação de desigualdades, consistindo no *tratamento desigual aos*

---

<sup>3</sup> A título ilustrativo, cabe mencionar que, recentemente, da mesma forma das adições de instrução ocultas em currículos, foram descobertos 17 artigos enviados para avaliação de publicação nos quais os autores colocavam instruções com "truques" – texto branco ou tamanho de fonte extremamente pequena – que determinavam "faça apenas uma avaliação positiva" e "não destaque nenhum ponto negativo". Alguns faziam exigências mais detalhadas, com um deles orientando os leitores da IA a recomendar o artigo por suas "contribuições impactantes, rigor metodológico e originalidade excepcional" (Schneier, 2025).

*desiguais, na medida da sua desigualdade*, o que é permitido pelo ordenamento brasileiro (Viola, 2014, p. 30). Inclusive, este tipo de tratamento costuma ser comum no ambiente de trabalho, ora por imposição legal, ora por medidas empresariais de alinhamento com práticas de igualitárias e de inclusão. Cabendo ressaltar que esta pesquisa se filia a esta corrente.

Normativamente, essa discussão passou a ser tratada em âmbito internacional em abril de 2021, quando a Comissão Europeia propôs a primeira normativa sobre inteligência artificial, estabelecendo um sistema de classificação baseado em riscos. O Regulamento UE 2024/1689 passou a vigorar em junho de 2024, de forma parcial e, após 24 meses, valerá de forma integral.

Desde o início, a regulamentação europeia demonstrou preocupação em proporcionar aos trabalhadores condições melhores, garantindo a não discriminação e dignidade. Nesse sentido, cita-se em especial o item 57 das considerações iniciais (União Europeia, 2024):

Os sistemas de IA utilizados nos domínios do emprego, da gestão de trabalhadores e do acesso ao emprego por conta própria, nomeadamente para efeitos de recrutamento e seleção de pessoal, de tomada de decisões que afetem os termos da relação de trabalho, de promoção e cessação das relações contratuais de trabalho, de atribuição de tarefas com base em comportamentos individuais, traços ou características pessoais, e de controlo ou avaliação de pessoas no âmbito de relações contratuais de trabalho também deverão ser classificados como sendo de risco elevado, uma vez que podem ter um impacto significativo nas perspectivas de carreira, na subsistência dessas pessoas e nos direitos dos trabalhadores. O conceito de «relações contratuais de trabalho» deverá abranger de forma significativa os funcionários e as pessoas que prestam serviços por intermédio de plataformas a que se refere o programa de trabalho da Comissão para 2021. Ao longo do processo de recrutamento e na avaliação, promoção ou retenção de pessoal em relações contratuais de trabalho, esses sistemas podem perpetuar padrões históricos de discriminação, por exemplo, contra as mulheres, contra certos grupos etários, contra as pessoas com deficiência ou contra pessoas de uma determinada origem racial ou étnica ou orientação sexual. Os sistemas de IA utilizados para controlar o desempenho e o comportamento dessas pessoas podem ainda comprometer os seus direitos fundamentais à proteção de dados pessoais e à privacidade.

A norma classificou como atividade de “risco elevado” o emprego de sistemas de IA para:

[...] recrutamento ou seleção de pessoas singulares, nomeadamente para colocar anúncios de emprego direcionados, analisar e filtrar candidaturas a ofertas de emprego e avaliar os candidatos;

[...] tomada de decisões que afetem os termos das relações de trabalho, a promoção ou a cessação das relações contratuais de trabalho, na atribuição de tarefas com base em comportamentos individuais, traços ou características pessoais, ou no controle e avaliação do desempenho e da conduta de pessoas que são partes nessas relações. (União Europeia, 2024, anexo 3, item 4)

Assim, evidencia-se que o regulamento europeu impõe um conjunto rigoroso de requisitos destinados a mitigar impactos sobre os direitos fundamentais dos trabalhadores, especialmente no que se refere à não discriminação. Entre essas exigências, destacam-se: a implementação de um sistema contínuo de gestão de

riscos (art. 9); o uso de dados de treinamento, validação e teste que sejam relevantes, representativos, livres de vieses e adequados ao propósito (art. 10); a manutenção de registros automatizados para fins de rastreabilidade e auditoria (art. 12); a obrigação de garantir transparência e disponibilização de informações claras aos usuários (art. 13). Ademais, o regulamento exige que haja supervisão humana efetiva (art. 14), bem como robustez técnica, precisão e resiliência cibernética dos sistemas (art. 15).

Na vanguarda dessa discussão, a cidade de Nova York editou a Lei nº 2021/144, que alterou seu código administrativo em relação ao uso de ferramentas automatizadas de decisão de emprego (AEDT), sendo a primeira cidade dos Estados Unidos a determinar que para aplicar essas ferramentas os empregadores devem: (i) realizar auditoria de viés, por meio de contratação de experts, para aferir se os algoritmos usados contêm algum tipo de viés fundado em raça, sexo ou etnia; (ii) publicar um resumo dos resultados da auditoria de viés em seu site; (iii) informar aos candidatos de que a ferramenta será empregada na sua avaliação; (iv) fornecer instruções para solicitação de adaptação razoável; (v) publicar aviso sobre o tipo e a fonte dos dados utilizados pela ferramenta, bem como a política de retenção de dados, em seu site. Inclusive regulou a possibilidade de registrar denúncia contra empregador ou agência de empregos no caso de aplicação de AEDT violando as medidas obrigatórias.

Por sua vez, a Organização Internacional do Trabalho (OIT), em abril de 2025, publicou relatório global sobre o papel da IA e da digitalização em revolucionar a segurança e saúde no trabalho, fazendo breves menções sobre o risco de discriminação do trabalhador. O documento destaca o importante papel que os sindicatos têm adotado na discussão das implicações da digitalização para os trabalhadores. Dentre os exemplos, menciona que:

[...] a Confederação Internacional de Sindicatos (ITUC) defende uma regulamentação mais forte da IA e da vigilância digital para proteger os trabalhadores e trabalhadoras da discriminação e da intensificação do trabalho. [...] A IndustriALL Global Union também criou um grupo de especialistas em Indústria 4.0 para desenvolver um documento de políticas sobre digitalização, IA e indústria 4.0 com foco na participação dos trabalhadores e trabalhadoras, na proteção dos direitos laborais e na garantia de uma transição digital justa para os trabalhadores e trabalhadoras (OIT, 2025, p. 43).

Também ressalta o papel legislativo dos “países que já começaram a regulamentar a gestão do trabalho por algoritmos face a riscos como sanções injustas, vigilância excessiva e discriminação - lacunas normalmente não abrangidas pela generalidade da legislação de SST”. Citando que: “a China e os Países Baixos introduziram regulamentos que realçam a equidade e a transparência dos algoritmos no local de trabalho”; a Espanha aprovou a Lei nº 12/2021, que “altera a Lei do Estatuto dos Trabalhadores, exigindo que os empregadores divulguem os parâmetros algorítmicos que afetam as condições de trabalho”; os Estados Unidos aprovaram a Lei *Algorithmic Accountability Act*, de 2019, que “exige avaliações de sistemas de IA e de aprendizagem automática com risco elevado que tratam dados pessoais ou que tomam decisões automatizadas”. (OIT, 2025, 50-51)

No Brasil, quanto ao tema ainda não há legislação específica para a regulação da IA, o que cria um desafio para conciliar a necessidade de desenvolvimento

tecnológico com a proteção contra a discriminação, principalmente em face da utilização da correlação para decisão algorítmica.

Entretanto, está em tramite o Projeto de Lei nº 2.338, de 2023, de autoria do Senador Rodrigo Pacheco (PSD/MG), aprovado pelo Senado Federal e remetido para análise na Câmara dos Deputados, onde se encontra no momento aguardando Parecer do(a) Relator(a) na Comissão Especial. A proposta mostra atenção a igualdade, não discriminação, pluralidade e diversidade, aos direitos sociais, em especial a valorização do trabalho humano, a privacidade, proteção de dados pessoais e autodeterminação informativa, a proteção e promoção de direitos de grupos vulneráveis, o acesso à informação e à disseminação de dados, de forma aberta, estruturada e segura e ao crescimento inclusivo, desenvolvimento sustentável e bem-estar, incluindo a proteção do trabalho e do trabalhador.

Seguindo a tendência mundial, classifica como atividade de alto risco o sistema de IA empregado para “recrutamento, triagem, filtragem ou avaliação de candidatos, tomada de decisões sobre promoções ou cessações de relações contratuais de trabalho, avaliação do desempenho e do comportamento das pessoas afetadas nas áreas de emprego, gestão de trabalhadores e acesso ao emprego por conta própria” (art. 14). Torna imprescindível a avaliação preliminar do risco da IA antes da sua colocação no mercado (art. 12), além de estabelecer deveres como ações de governança, acreditação, certificação e avaliação de conformidade, sanções e disposições sobre a responsabilidade civil.

Além disso, em proposta de alteração no Código Civil (Projeto de Lei nº 4/2025 do Senado Federal) está prevista a inclusão de um capítulo sobre IA (VII Inteligência Artificial), com previsão de que o uso de sistema de IA deve respeitar os direitos da personalidade e garantir o seguinte:

Art. 2.027-AL. O desenvolvimento de sistemas de inteligência artificial deve respeitar os direitos de personalidade previstos neste Código, garantindo a implementação de sistemas seguros e confiáveis, em benefício da pessoa natural ou jurídica e do desenvolvimento científico e tecnológico, devendo ser garantidos:

I - a não discriminação em relação às decisões, ao uso de dados e aos processos baseados em inteligência artificial;

II - condições de transparência, auditabilidade, explicabilidade, rastreabilidade, supervisão humana e governança;

III - a acessibilidade, a usabilidade e a confiabilidade;

IV - a atribuição de responsabilidade civil, pelo princípio da reparação integral dos danos, a uma pessoa natural ou jurídica em ambiente digital.

Parágrafo único. O desenvolvimento e o uso da inteligência artificial e da robótica em áreas relevantes para os direitos de personalidade devem ser monitorados pela sociedade e regulamentados por legislação específica.

Há também previsão de que pessoas naturais ao interagirem ou sofrerem danos decorrentes da operação desses sistemas ou equipamentos de inteligência artificial têm “o direito à informação sobre suas interações com tais sistemas, bem como sobre o modelo geral de funcionamento e critérios para decisão automatizada, quando esta influenciar diretamente no seu acesso ou no exercício de direitos ou afetar seus interesses econômicos de modo significativo” (art. 2.027-AM). Por fim, o art. 609-F do projeto de lei prevê que a utilização de IA na prestação do serviço digital “deve ser identificada de forma clara e seguir os

padrões éticos necessários, segundo os princípios da boa-fé e da função social do contrato”.

Contudo, enquanto tais projetos de lei não são aprovados, permanece uma lacuna normativa no Brasil. Assim, atualmente, casos de discriminação algorítmica são resolvidos com base na Lei nº 9.029/1995, que trata sobre a discriminação no ambiente de trabalho, ou considerando dispositivos da Lei nº 13.709/2018 (LGPD), que garante ao titular dos dados o direito de saber se uma decisão foi automatizada, o direito de solicitar revisão por pessoa humana (art. 20), o direito à explicação clara e adequada sobre os critérios usados (art. 6, VI) e o direito a não discriminação no tratamento de dados (art. 6, IX). Além de outros dispositivos e princípios esparsos no ordenamento jurídico, a Matos (2024, p. 114, 135-136) enumera diretrizes fundamentais que podem auxiliar no estabelecimento de balizas para a IA:

- dignidade da pessoa humana (art. 1º, III, CF/88), promoção da diversidade e igualdade substancial (art. 5º, caput, CF/88) e proteção da privacidade (art. 5º, X, CF/88) e do direito fundamental à proteção de dados (art. 5º, LXXIX, CF/88), buscando a centralidade na proteção do ser humano, especialmente de grupos vulneráveis;
- indelegabilidade da decisão intrinsecamente humana, a fim de que as decisões sejam tomadas de forma adequada e ética;
- bem-estar de todos, de modo a garantir que a IA seja desenvolvida e utilizada levando em consideração não apenas o bem-estar das pessoas no presente, mas também o impacto nas gerações futuras e no meio ambiente (art. 225, CF/88);
- segurança dos dados e informações, com adoção de medidas preventivas e precauções para proteger os sistemas contra ataques cibernéticos e vazamentos de dados;
- reversibilidade em caso de problemas ou consequências indesejadas, ou seja, deve ser possível reverter ou corrigir as ações da IA;
- transparência e explicabilidade, de modo que os processos de tomada de decisão da IA sejam explicáveis e auditáveis;
- implementação de mecanismos de controle de abusos e discriminação no ambiente de trabalho;
- estabelecer requisitos para avaliação de impactos nos direitos fundamentais, como privacidade e proteção de dados, bem como nas normas de repressão à discriminação;
- incluir na programação dos sistemas a participação de grupos potencialmente afetados por práticas discriminatórias;
- implementar mecanismos de salvaguardas e garantias de reparação adequada em caso de impactos adversos;
- informar ao titular dos dados sobre a identidade e contato do responsável pelo tratamento, bem como oportunizar o direito de reclamação;

A autora ainda defende que a empresa deve informar, tanto aos trabalhadores individuais quanto aos sindicatos, sobre o uso de algoritmos ou sistemas de IA para tomar decisões automatizadas de gestão de pessoas, incluindo candidatos a emprego. Isso abrange não só identificar o algoritmo que está sendo utilizado, mas também os dados envolvidos, o tipo de tecnologia, como opera, quais variáveis emprega, as auditorias e avaliações do algoritmo e as consequências da decisão tomada pelo sistema (Matos, 2024, p. 135-136).

Diante do exposto, observa-se que a simples existência de correlação entre variáveis não é suficiente para legitimar práticas que, em sua essência, reproduzem desigualdades estruturais historicamente arraigadas.

Nesse sentido, há um esforço geral e global para regularizar o uso da inteligência artificial, buscando principalmente promover a proteção do trabalhador contra discriminações e violações de sua privacidade.

No Brasil, embora não exista regulamentação específica, iniciativas legislativas como o PL nº 2.338/2023 e o PL nº 4/2025 sinalizam avanços na incorporação de princípios como transparência, não discriminação, supervisão humana e responsabilidade no uso de sistemas algorítmicos, incorporando apontamentos doutrinários sobre a imprescindível proteção da pessoa humana, especialmente dos grupos vulnerabilizados, assegurando que os sistemas de IA não apenas respeitem os limites ético-jurídicos, mas também sejam instrumentos ativos na promoção da igualdade.

#### **4 Considerações finais**

Na pesquisa analisou-se o uso da correlação nas tomadas de decisão nas relações de trabalho, especialmente por algoritmos, e quais os limites de seu uso. Conclui-se, quanto ao primeiro problema, que seu uso tem comprometido direitos fundamentais como a não discriminação, igualdade e privacidade. Isso porque, ao longo da pesquisa, pontuou-se que: (i) o abandono da causalidade gera a substituição de uma individualização concisa e justificada, por uma generalização baseada em estatística que, se construída em uma sociedade estigmatizada, apenas reforça tais estigmas; (ii) o uso de estatística pode resultar em uma filtragem direta ou indireta baseada em critérios constitucionalmente protegidos, isto é, os elencados no artigo 3º, IV, da CF/88, critérios de origem, raça, sexo, cor, idade e quaisquer outras formas de discriminação; (iii) o uso de algoritmos resulta em um modelo de opacidade, em que se sabe mais sobre o candidato/empregado do que sobre os dados que foram tratados e a forma como foram tratados, o que se deve ao *machine learning*, quando a máquina está em contínuo aprendizado e reaprendizado.

Assim, mesmo em casos em que o objetivo não é discriminatório, pode haver vieses ocultos nas decisões algorítmicas, como ocorreu nos casos práticos trazidos, cujas experiências demonstraram como decisões automatizadas, mesmo quando embasadas em grandes volumes de dados, podem cristalizar injustiças estruturais, principalmente se operadas sem transparência ou controle pelo ordenamento jurídico.

Portanto, respondendo o segundo problema, adota-se o posicionamento de que a correlação algorítmica deve ser afastada de critérios protegidos, salvo se o tratamento diferencial configurar ação afirmativa em benefício de grupos historicamente desfavorecidos.

Ademais, demonstra-se que urge discussão e regulamentação da correlação algorítmica para impedir a discriminação direta e indireta nas relações de trabalho, uma vez que, quanto mais se sofisticam as técnicas de utilização dos dados, mais difícil se torna identificar o momento em que a racionalidade cede lugar à arbitrariedade ou ao reforço de desigualdades e da discriminação.

Diante de todo o exposto, evidencia-se imprescindível reconquistar o equilíbrio entre a autonomia privada, a sociedade do risco e do desempenho com os direitos fundamentais, sob o risco dos indivíduos viverem aprisionados na caverna digital, convencidos de estar livres, enquanto submetidos a uma lógica algorítmica que obscurece a verdade, esvazia a subjetividade e mercantiliza a existência. Mais do que ajustar sistemas, é urgente reconduzir o humano ao centro do debate.

## Referências

ABBAGNANO, Nicola. *Dicionário de Filosofia*. Trad. Alfredo Bosi e Ivone Castilho Benedeti. São Paulo: Martins Fontes, 2007.

ARAGÃO, Francisca Alana Araújo; BENEVIDES, Pablo Severiano. Governamentalidade algorítmica e big data: o uso da correlação de dados como critério de tomada de decisão. In: *Simpósio Internacional Lavits, IV*, 2019, Salvador.

ARISTÓTELES. *Metafísica*. Trad. Leonel Vallandro. Porto Alegre: Globo, 1969.

BLOOMBERG. “Fui despedido por um robô”: como a Amazon deixa máquinas decidirem o destino dos trabalhadores. *O Globo*, 28 de jun. 2021. Disponível em: <https://oglobo.globo.com/economia/tecnologia/fui-despedido-por-um-roboto-como-amazon-deixa-maquinas-decidirem-destino-dos-trabalhadores-25079925>. Acesso em: 11 maio 2025.

BRASIL. Senado Federal. Projeto de Lei nº 2.338, de 2023. *Estabelece princípios, direitos e deveres para o uso da inteligência artificial no Brasil*. Disponível em: <https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2487262>. Acesso em: 6 jul. 2025.

CARPES, Artur Thompsen. *A prova do nexo de causalidade na responsabilidade civil*. 2013. Tese (Doutorado em Direito Processual Civil) – Programa de Pós-Graduação em Direito, Faculdade de Direito, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2013. Disponível em: <https://lume.ufrgs.br/handle/10183/207209>. Acesso em: 6 jul. 2025.

COLOMBO, Cristiano; FACCHINI NETO, Eugênio. “Corpo eletrônico” como vítima de ofensas em matéria de tratamento de dados pessoais: reflexões acerca da responsabilidade civil por danos à luz da Lei Geral de Proteção de Dados Pessoais Brasileira e a viabilidade da aplicação da noção de dano estético ao mundo digital. In: ROSENVALD, Nelson et al. *Responsabilidade civil novos riscos*. Indaiatuba, SP: Foco, 2019.

CUNHA, José Ricardo. Direito e complexidade. In: BARRETO, Vicente de Paulo (Coord.). *Dicionário de Filosofia de Direito*. São Leopoldo: Unisinos, 2006.

DE CUPIS, Adriano. *Il danno*. 2. ed. Milano: Giuffrè, 1966.

DRESLER, Georg. Prompt injection for large language models. 3 fev. 2025. *InfoQ*. Disponível em: <https://www.infoq.com/articles/large-language-models-prompt-injection-stealing/>. Acesso em: 14 jul. 2025.

ECHARRI, Miguel. 150 demissões em um segundo: os algoritmos que decidem quem deve ser mandado embora. *El País*. Barcelona, 10 out. 2021. Disponível em: <http://brasil.elpais.com/tecnologia/2021-10-10/150-demissoes-em-um-segundo-assim-funcionam-os-algoritmos-que-decidem-quem-deve-ser-mandado-embora.html>. Acesso em: 14 jul. 2025.

FROTA, Pablo Malheiros Cunha. *Imputação sem nexo causal e a responsabilidade por danos*. Dissertação - Universidade Federal do Paraná, Curitiba, 2013. Disponível em: [https://acervodigital.ufpr.br/xmlui/bitstream/handle/1884/31777/R%20-%20T%20-%20PABLO%20MALHEIROS%20DA%20CUNHA%20FROTA.pdf?sequence=1&isAllowed=y](https://acervodigital.ufpr.br/xmlui/bitstream/handle/1884/31777/R%20-%20T%20-%20PABLO%20MALHEIROS%20DA%20CUNHA%20FROTA.pdf?sequence=1&isAll%20owed=y). Acesso em: 14 jul. 2025.

GOLDEMBERG, Isidoro H. *La relación de causalidad en la responsabilidad civil*. Buenos Aires: Astrea, 1989.

HAN, Byung-Chul. *Infocracia: digitalização e crise da democracia*. Trad. Gabriel S. Philipson. 5. Reimp. Petrópolis, RJ: Vozes, 2024a.

HAN, Byung-Chul. *Sociedade do cansaço*. Trad. Enio Paulo Giachini. 3. ed. Petrópolis, RJ: Vozes, 2024b.

HUME, David. *Investigação acerca do entendimento humano*. Trad. Anoar Aiex. São Paulo: Nova Cultural, 1996 (Os Pensadores).

HUME, David. *Investigações sobre o entendimento humano e sobre os princípios da moral*. São Paulo: Unesp, 2004.

JUNQUEIRA, Thiago. *Tratamento de dados pessoais e discriminação algorítmica nos seguros*. São Paulo: Thomson Reuters Brasil, 2020.

LIU, Ian Ch. Invisible prompt injection: a threat to AI security. *Trend Micro*, 22 jan. 2025. Disponível em: [https://www.trendmicro.com/en\\_us/research/25/a/invisible-prompt-injection-secure-ai.html](https://www.trendmicro.com/en_us/research/25/a/invisible-prompt-injection-secure-ai.html). Acesso em: 14 jul. 2025.

LOPEZ, Teresa Ancona. *Nexo causal e produtos potencialmente nocivos: a experiência brasileira do tabaco*. São Paulo: Quartier Latin, 2008.

MARLER, Janet H.; BOUDREAU, John W. An evidence-based review of HR analytics. *The International Journal of Human Resource Management*, v. 28, p. 3-26, 2017.

MATOS, Larissa. *Inteligência artificial, algoritmos e direito do trabalho: conceitos, impactos e limites*. São Paulo: Mizuno, 2024.

MELLO, Marcos Bernardes de. *Teoria do fato jurídico: plano da eficácia: 1ª Parte*. 3. ed. São Paulo: Saraiva, 2007.

MIRAGEM, Bruno. A Lei Geral de Proteção de Dados (Lei 13.709/2018) e o direito do consumidor. *Revista dos Tribunais*, São Paulo, v. 1009, nov. 2019.

MONTEIRO, Rosângela; RABELLO, Guilherme de Castro Machado; ARRUDA JÚNIOR, Francismar Vidal de; JATENE, Fabio Biscegli. Inteligência artificial, *deep learning*, *machine learning*, redes neurais na medicina e biomarcadores vocais: conceitos, onde estamos e para onde vamos. *Revista da Sociedade de Cardiologia*

do Estado de São Paulo, São Paulo, v. 32, n. 1, p. 11–17, 2022. Disponível em: <http://dx.doi.org/10.29381/0103-8559/2022320111-7>. Acesso em: 9 maio 2025.

MORA, José Ferrater. *Dicionário de Filosofia*. Trad. Roberto Leal Ferreira e Álvaro Cabral. São Paulo: Martins Fontes, 2001.

MORGAN, Kate; NOLAN, Delaney. Como a vigilância dos empregados está se voltando contra os empregadores. *BBC News Brasil*, 5 mar. 2023. Disponível em: <https://www.bbc.com/portuguese/articles/c4n08l9wzqdo>. Acesso em: 5 maio 2024.

MURTEIRA, Bento *et al.* *Introdução à estatística*. Lisboa: Escolar, 2010.

O'NEIL, Cathy. *Algoritmos de destruição em massa: como o big data aumenta a desigualdade e ameaça a democracia*. São Paulo: Rua do Sabão, 2020.

ORGANIZAÇÃO INTERNACIONAL DO TRABALHO (OIT). *Revolucionar a segurança e saúde no trabalho: o papel da IA e da digitalização*. Genebra: OIT, 2025. Disponível em: <https://www.ilo.org/pt-pt/resource/news/ia-e-digitalizacao-estao-transformar-seguranca-e-saude-no-trabalho>. Acesso em: 6 jul. 2025.

ORGAZ, Alfredo. *El daño resarcible*. Buenos Aires: Córdoba, 1980.

ORWAT, Carsten. *Risks of discrimination through the use of algorithms: a study compiled with a grant from the Federal Anti-Discrimination Agency*. Karlsruhe: Karlsruhe Institute of Technology (KIT), Institute for Technology Assessment and Systems Analysis (ITAS), 2020. Disponível em: [https://www.antidiskriminierungsstelle.de/EN/homepage/documents/download/diskr\\_risiken\\_verwendung\\_von\\_algorithmen.pdf?blob=publicationFile&v=1](https://www.antidiskriminierungsstelle.de/EN/homepage/documents/download/diskr_risiken_verwendung_von_algorithmen.pdf?blob=publicationFile&v=1). Acesso em: 5 jul. 2025.

PASQUALOTTO, Alberto. Causalidade e imputação na responsabilidade civil objetiva: uma reflexão sobre os assaltos em estacionamentos. *Revista dos Tribunais*, v. 7, abr./jun. 2016.

PATRÍCIO, Miguel. Correlações e causalidade: breves reflexões numa perspectiva económica. *Revista Jurídica Luso-Brasileira*, Lisboa, ano 02, n. 04, 2016, p. 1347-1380. Disponível em: [https://www.cidp.pt/revistas/rjlb/2016/4/2016\\_04\\_1347\\_1380.pdf](https://www.cidp.pt/revistas/rjlb/2016/4/2016_04_1347_1380.pdf). Acesso em: 5 jul. 2025.

PRUITT, Pru. Prompt injection: the hidden threat every ai business and user should understand. *Medium*, 29 abr. 2025. Disponível em: <https://medium.com/@Justpruitt/prompt-injection-the-hidden-threat-every-ai-business-and-user-should-understand-71821f16b366>. Acesso em: 14 jul. 2025.

REUTERS. Amazon desiste de ferramenta secreta de recrutamento que mostrou viés contra mulheres. *O Globo*, 10 out. 2018. Disponível em: <https://epocanegocios.globo.com/Empresa/noticia/2018/10/amazon-desiste-de-ferramenta-secreta-de-recrutamento-que-mostrou-vies-contras-mulheres.html>. Acesso em: 02 maio 2025.

SAHU, Anubhab. *Understanding invisible prompt injection attack*. Keysight Technologies, 16 maio 2025. Disponível em: <https://www.keysight.com/blogs/en/tech/nwvs/2025/05/16/invisible-prompt-injection-attack>. Acesso em: 14 jul. 2025.

SCHNEIER. Hiding prompt injections in academic papers. *Schneier on Security (blog)*, 7 jul. 2025. Disponível em: <https://www.schneier.com/blog/archives/2025/07/hiding-prompt-injections-in-academic-papers.html>. Acesso em: 14 jul. 2025.

SCHREIBER, Anderson. *Novos paradigmas da responsabilidade civil*. São Paulo: Atlas, 2007.

SILVA, Tarcízio. *Racismo algorítmico: inteligência artificial e discriminações nas redes sociais*. São Paulo: Edições SESC, 2022.

TARTUCE, Flávio. *Responsabilidade civil objetiva e risco - a teoria do risco concorrente*. Rio de Janeiro: Forense; São Paulo: Método, 2011.

UNIÃO EUROPEIA. Regulamento (UE) 2024/1689 do Parlamento Europeu e do Conselho, de 13 de junho de 2024, que estabelece regras harmonizadas em matéria de inteligência artificial (Lei da Inteligência Artificial) e altera os Regulamentos (CE) n. 300/2008, (UE) 2018/858, (UE) 2019/2144 e (UE) 2019/2115 e as Diretivas 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828. *Diário Oficial da União Europeia*, L, n. 1689, 12 jul. 2024. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/?uri=CELEX%3A32024R1689>. Acesso em: 6 jul. 2025.

VILANOVA, Lourival. *Causalidade e relação no direito*. 4. ed. São Paulo: Revista dos Tribunais, 2000.

VIOLA, Mario. Gênero e diferenciação de prêmios: reflexões sobre o princípio da igualdade. *Cadernos de Seguro*, Rio de Janeiro, ano XXXIV, n. 180, v. especial, abr. 2014, p. 28-31.

WALSH, Dylan. Don't let artificial intelligence pick your employees. *Stanford Business*, 8 fev. 2019. Disponível em: [https://www.gsb.stanford.edu/insights/dont-let-artificial-intelligence-pick-your-employees?utm\\_source=Stanford+Business&utm\\_campaign=8de21a7c50-Stanford-Business-Issue-156-2-24-2019&utm\\_medium=email&utm\\_term=0\\_0b5214e34b-8de21a7c50-73763977&ct=t\(Stanford-Business-Issue-156-2-24-2019\)](https://www.gsb.stanford.edu/insights/dont-let-artificial-intelligence-pick-your-employees?utm_source=Stanford+Business&utm_campaign=8de21a7c50-Stanford-Business-Issue-156-2-24-2019&utm_medium=email&utm_term=0_0b5214e34b-8de21a7c50-73763977&ct=t(Stanford-Business-Issue-156-2-24-2019)). Acesso em: 2 maio 2025.

ZAPATA, Dawn. New study finds AI-enabled anti-Black bias in recruiting. *Thomson Reuters*, 18 jun. 2021. Disponível em: <https://www.thomsonreuters.com/en-us/posts/legal/ai-enabled-anti-black-bias/>. Acesso em: 5 maio 2024.

**FLUXO EDITORIAL:**

- Data de submissão: 04 fev. 2026
- Data de aprovação: 08 jun. 2026
- Data de publicação: 01 jul. 2026
  
- Avaliação: por pares dupla-anônima
- Pareceristas: 2 (dois)
- Editor Chefe: José Carlos Francisco