

A PROTEÇÃO INTERNACIONAL DOS DIREITOS HUMANOS DIANTE DA INTELIGÊNCIA ARTIFICIAL: LIMITES E DESAFIOS PARA OS SISTEMAS INTERNACIONAIS DE PROTEÇÃO

**THE INTERNATIONAL PROTECTION OF HUMAN RIGHTS IN THE FACE OF
ARTIFICIAL INTELLIGENCE: LIMITS AND CHALLENGES FOR
INTERNATIONAL PROTECTION SYSTEMS**

Ciências Sociais Aplicadas • 07/10/2026

REGISTRO DOI: [10.70773/revistatopicos/791356437](https://doi.org/10.70773/revistatopicos/791356437)

Camila Dantas Rocha de Araújo¹

RESUMO

O desenvolvimento e a expansão da Inteligência Artificial têm produzido mudanças relevantes nas relações sociais, econômicas e institucionais, ao mesmo tempo em que ampliam situações capazes de afetar direitos reconhecidos internacionalmente. Sistemas automatizados empregados em serviços públicos, segurança, saúde, trabalho, crédito, plataformas digitais e processos decisórios podem interferir na privacidade, na igualdade, na liberdade de expressão, no acesso à informação, no devido processo e no direito à reparação. Este estudo teve como objetivo analisar os limites e os desafios enfrentados pelos sistemas internacionais de proteção dos Direitos Humanos diante do desenvolvimento e do uso da Inteligência Artificial. Realizou-se pesquisa bibliográfica, de abordagem qualitativa e caráter exploratório e descritivo, complementada pela análise de documentos produzidos pela Organização das Nações Unidas, UNESCO, Conselho da Europa, União Europeia e Comissão Interamericana de Direitos Humanos. Os resultados indicaram que os instrumentos internacionais vigentes oferecem fundamentos jurídicos aplicáveis aos impactos da Inteligência Artificial, porém apresentam dificuldades relacionadas à transparência algorítmica, à identificação de responsabilidades, à atuação transnacional das empresas e ao acesso a mecanismos efetivos de reparação. Verificou-se ainda uma evolução normativa relevante, especialmente com a adoção da Convenção Quadro do Conselho da Europa sobre Inteligência Artificial. Concluiu-se que a proteção internacional depende menos da criação de novos direitos e mais da capacidade de aplicar direitos já reconhecidos às novas formas de decisão e controle mediadas por sistemas automatizados.

Palavras-chave: Direitos Humanos; Inteligência Artificial; Proteção Internacional; Responsabilidade Algorítmica; Sistemas Internacionais.

ABSTRACT

The development and expansion of Artificial Intelligence have produced significant changes in social, economic and institutional relations while increasing situations capable of affecting internationally recognized rights. Automated systems used in public services, security, health care, employment, credit, digital platforms and decision making may interfere with privacy, equality, freedom of expression, access to information, due process and the right to an effective remedy. This study aimed to analyze the limits and challenges faced by international Human Rights protection systems in response to the development and use of Artificial Intelligence. Bibliographical research was conducted through a qualitative, exploratory and descriptive approach, complemented by an analysis of documents produced by the United Nations, UNESCO, the Council of Europe, the European Union and the Inter American Commission on Human Rights. The results indicated that existing international instruments provide legal foundations applicable to the impacts of Artificial Intelligence, although difficulties remain regarding algorithmic transparency, attribution of responsibility, the transnational activities of companies and access to effective remedies. Relevant normative development was also identified, particularly with the adoption of the Council of Europe Framework Convention on Artificial Intelligence. The study concluded that international protection depends less on the creation of new rights than on the capacity to apply already recognized rights to new forms of decision making and control mediated by automated systems.

Keywords: Human Rights; Artificial Intelligence; International Protection; Algorithmic Accountability; International Protection Systems.

INTRODUÇÃO

A consolidação da Inteligência Artificial como tecnologia aplicada a atividades públicas e privadas alterou a maneira como informações são produzidas, classificadas e utilizadas na tomada de decisões. Sistemas capazes de reconhecer padrões, produzir previsões, classificar indivíduos e gerar conteúdos passaram a integrar áreas que possuem relação direta com o exercício de direitos. O uso dessas ferramentas deixou de permanecer restrito à automação de tarefas técnicas e alcançou procedimentos relacionados à concessão de benefícios públicos, recrutamento profissional, análise de crédito, segurança, identificação biométrica, saúde, produção e circulação de informações e funcionamento de plataformas digitais. A mudança possui relevância jurídica porque parte das decisões que interferem na vida das pessoas passou a depender de processos computacionais cujo funcionamento nem sempre pode ser conhecido por aqueles que são atingidos por seus resultados.

Os Direitos Humanos foram estruturados internacionalmente a partir da compreensão de que a dignidade da pessoa deve estabelecer limites ao exercício do poder público e orientar a atuação institucional. A Declaração Universal dos Direitos Humanos, adotada em 1948, consolidou direitos relacionados à igualdade, à liberdade, à privacidade, ao acesso à justiça e à proteção contra formas arbitrárias de interferência na vida individual. Posteriormente, tratados internacionais transformaram parte desses compromissos em obrigações jurídicas assumidas pelos Estados. Embora esses documentos tenham sido elaborados em períodos muito anteriores à atual capacidade dos sistemas de Inteligência Artificial, seus fundamentos continuam aplicáveis quando as formas de exercício do poder são modificadas pela tecnologia.

A dificuldade atual não decorre, portanto, apenas da ausência de normas especificamente destinadas à Inteligência Artificial. O problema envolve a interpretação e a aplicação de direitos previamente reconhecidos em situações caracterizadas por grande capacidade de processamento de dados, utilização de modelos complexos e participação simultânea de agentes públicos e privados. McGregor, Murray e Ng (2019) sustentam que o Direito Internacional dos Direitos Humanos já fornece elementos relevantes para a responsabilização algorítmica porque estabelece parâmetros para identificação de danos, prevenção de violações e proteção das pessoas afetadas. Entretanto, a aplicação desse conjunto normativo enfrenta dificuldades quando decisões são formadas por processos pouco transparentes ou quando diferentes empresas participam das etapas de desenvolvimento, fornecimento e utilização de uma mesma tecnologia.

O direito à privacidade representa um dos pontos mais evidentes dessa relação. Sistemas de Inteligência Artificial dependem frequentemente de grandes quantidades de informações para treinamento, identificação de padrões e geração de resultados. Dados originalmente coletados para determinada finalidade podem ser combinados com outras fontes, permitindo a elaboração de perfis e inferências que não eram conhecidas pelo próprio titular das informações. Wachter e Mittelstadt (2019) apontam que a análise de grandes bases de dados pode produzir inferências sobre comportamentos, preferências e características pessoais cuja validade e utilização são de difícil contestação. O problema ultrapassa a coleta convencional de dados e alcança a capacidade de produzir conhecimento novo sobre indivíduos a partir da combinação automatizada de informações.

A preocupação também se estende ao direito à igualdade. A utilização de modelos estatísticos pode reproduzir desigualdades contidas nos dados utilizados em seu treinamento. Barocas e Selbst (2016) demonstram que procedimentos computacionais aparentemente neutros são capazes de reproduzir padrões históricos de exclusão quando os dados refletem práticas discriminatórias existentes na sociedade. Uma decisão automatizada pode, dessa forma, gerar tratamento desigual mesmo sem a inclusão explícita de critérios relacionados a raça, gênero, deficiência ou origem social. Variáveis aparentemente neutras podem funcionar como indicadores indiretos dessas características, dificultando a identificação jurídica do tratamento discriminatório.

Mittelstadt et al. (2016) observam que a delegação de decisões a algoritmos também modifica a discussão sobre responsabilidade. Uma decisão humana tradicional permite, com maior facilidade, identificar a pessoa ou a instituição responsável pela escolha realizada. Sistemas de Inteligência Artificial podem envolver desenvolvedores, fornecedores de infraestrutura, empresas responsáveis pelos dados, organizações contratantes e profissionais que utilizam os resultados. Essa distribuição funcional não elimina a responsabilidade, mas torna mais difícil determinar quem possuía controle suficiente para prevenir o dano e quem deve responder quando um direito é violado.

A incorporação de sistemas automatizados pela Administração Pública torna essa questão especialmente importante. A utilização de tecnologias para selecionar beneficiários, identificar fraudes, estabelecer níveis de risco ou direcionar atividades de fiscalização pode aumentar a capacidade administrativa, mas também cria novas formas de interferência estatal. Rachovitsa e Johann (2022), ao

analisarem o caso do sistema holandês SyRI, demonstraram que o uso de ferramentas algorítmicas em políticas sociais pode afetar direitos quando existe opacidade sobre os critérios utilizados e quando determinados grupos são submetidos de maneira desproporcional a mecanismos de vigilância e classificação.

A proteção internacional enfrenta ainda o problema da territorialidade. Empresas responsáveis pelo desenvolvimento de sistemas de Inteligência Artificial podem possuir sede em determinado país, realizar processamento de dados em outro e disponibilizar seus produtos para usuários localizados em várias jurisdições. Esse modelo reduz a capacidade de respostas exclusivamente nacionais. Lane (2022) identifica que a multiplicação de iniciativas regulatórias em diferentes níveis contribui para esclarecer determinados padrões de Direitos Humanos, mas ainda permanecem incertezas sobre a extensão das obrigações atribuídas aos diversos agentes envolvidos na criação e na aplicação da Inteligência Artificial.

O sistema universal das Nações Unidas passou a dedicar maior atenção ao tema. O Alto Comissariado das Nações Unidas para os Direitos Humanos (2021) analisou o impacto da Inteligência Artificial sobre a privacidade e direitos relacionados, destacando problemas associados à vigilância, ao reconhecimento biométrico, à criação de perfis e às decisões automatizadas. A análise partiu da compreensão de que tecnologias capazes de ampliar significativamente o conhecimento sobre a vida das pessoas podem alterar a relação de poder entre indivíduos, empresas e governos. A questão central deixou de ser somente a proteção de informações isoladas e passou a envolver o controle sobre processos capazes de interpretar essas informações e produzir consequências jurídicas ou sociais.

A UNESCO (2021) também adotou uma Recomendação sobre a Ética da Inteligência Artificial, estruturada sobre a proteção da dignidade humana, dos Direitos Humanos, da diversidade e da autonomia. Embora a Recomendação não possua a mesma força jurídica de um tratado internacional, sua aprovação representou avanço relevante na construção de referências compartilhadas. O documento incorporou temas como proporcionalidade, segurança, equidade, não discriminação, supervisão humana, transparência e responsabilidade, aproximando o debate ético sobre a tecnologia do conjunto de direitos reconhecidos internacionalmente.

Em 2024, a Assembleia Geral das Nações Unidas aprovou a Resolução A/RES/78/265, voltada à promoção de sistemas de Inteligência Artificial seguros, protegidos e confiáveis. A Organização das Nações Unidas (2024) reconheceu a necessidade de reduzir desigualdades digitais e promover tecnologias compatíveis com os Direitos Humanos e com os objetivos de desenvolvimento sustentável. A resolução não criou um regime internacional de responsabilidade, porém reforçou a compreensão de que a governança da Inteligência Artificial não pode ser separada das obrigações internacionais já existentes.

A evolução mais expressiva no campo convencional ocorreu no Conselho da Europa. Em 2024, foi adotada a Convenção Quadro sobre Inteligência Artificial, Direitos Humanos, Democracia e Estado de Direito. O Conselho da Europa (2024) estabeleceu um instrumento juridicamente vinculante destinado a assegurar que atividades realizadas durante o ciclo de vida dos sistemas de Inteligência Artificial permaneçam compatíveis com Direitos Humanos, democracia e Estado de Direito. A Convenção incorporou princípios de dignidade, autonomia, igualdade, transparência,

supervisão, responsabilidade, privacidade e confiabilidade, além de prever avaliação de riscos e impactos.

No âmbito europeu, a regulamentação também avançou com o Regulamento (UE) 2024/1689, conhecido como Regulamento de Inteligência Artificial. A União Europeia (2024) estruturou obrigações de acordo com os riscos associados às diferentes utilizações da tecnologia, estabelecendo proibições para determinadas práticas e exigências mais rigorosas para sistemas classificados como de alto risco. Embora seja um instrumento regional e possua objetivos relacionados ao funcionamento do mercado europeu, suas regras interagem diretamente com direitos fundamentais e influenciam debates regulatórios fora da própria União Europeia.

No sistema interamericano, a consolidação de um marco específico encontra-se em estágio distinto. A Convenção Americana sobre Direitos Humanos contém garantias suficientemente amplas para alcançar situações relacionadas à Inteligência Artificial, incluindo proteção da honra e da dignidade, igualdade perante a lei, garantias judiciais, liberdade de pensamento e expressão e proteção judicial. A Comissão Interamericana de Direitos Humanos passou a acompanhar de forma crescente os efeitos das tecnologias digitais e da Inteligência Artificial, especialmente em relação à liberdade de expressão, à manipulação informacional, à discriminação, à privacidade e à proteção de grupos historicamente vulneráveis. Ainda assim, permanece espaço para maior sistematização de parâmetros regionais destinados às decisões automatizadas.

O problema desta pesquisa foi formulado da seguinte maneira: em que medida os sistemas internacionais de proteção dos Direitos Humanos possuem instrumentos capazes de responder às violações

e aos riscos decorrentes do desenvolvimento e da utilização da Inteligência Artificial? A pergunta parte do reconhecimento de que a existência formal de direitos não garante, por si, proteção efetiva quando novas tecnologias modificam os mecanismos pelos quais decisões são tomadas, informações são utilizadas e danos são produzidos.

O objetivo geral consistiu em analisar os limites e os desafios enfrentados pelos sistemas internacionais de proteção dos Direitos Humanos diante da Inteligência Artificial. Como objetivos específicos, buscou-se examinar os principais direitos suscetíveis de impacto por decisões algorítmicas; identificar os mecanismos desenvolvidos no sistema universal e nos sistemas regionais; analisar dificuldades de transparência, responsabilização e reparação; e discutir a possibilidade de aplicação dos instrumentos internacionais existentes às novas formas de atuação tecnológica.

A relevância do estudo decorre da necessidade de evitar que a inovação tecnológica seja tratada como espaço separado do Direito. A Inteligência Artificial modifica instrumentos e procedimentos, mas não afasta as obrigações jurídicas dos Estados nem a responsabilidade das empresas pelo respeito aos Direitos Humanos. Quanto maior a participação dos sistemas automatizados em decisões socialmente relevantes, maior se torna a necessidade de assegurar que eficiência, inovação e capacidade computacional não substituam garantias construídas historicamente para proteger a pessoa diante do exercício do poder.

METODOLOGIA

A pesquisa foi desenvolvida por meio de pesquisa bibliográfica, com abordagem qualitativa e caráter exploratório e descritivo. A escolha desse método decorreu da natureza jurídica do problema investigado, que exigiu o exame articulado da produção científica sobre Inteligência Artificial e Direitos Humanos e dos instrumentos internacionais aplicáveis à matéria. A análise não buscou mensurar quantitativamente o uso de sistemas automatizados, mas identificar categorias jurídicas, dificuldades de proteção e respostas institucionais presentes na literatura e nos sistemas internacionais.

O levantamento bibliográfico considerou artigos científicos, capítulos de livros e estudos jurídicos relacionados a Inteligência Artificial, decisões automatizadas, privacidade, discriminação algorítmica, responsabilidade e Direito Internacional dos Direitos Humanos. Foram priorizadas publicações com relação direta ao problema de pesquisa e trabalhos que permitissem compreender as consequências jurídicas da utilização de algoritmos em contextos capazes de afetar direitos fundamentais. Entre os autores utilizados estão Barocas e Selbst (2016), Mittelstadt et al. (2016), McGregor, Murray e Ng (2019), Wachter e Mittelstadt (2019), Lane (2022), Rachovitsa e Johann (2022), Krupiy e Scheinin (2023), Kamble (2024), Matar e Murray (2025) e McGregor (2026).

A seleção bibliográfica privilegiou estudos publicados em periódicos científicos e obras acadêmicas reconhecidas, com atenção especial às áreas de Direito Internacional, Direitos Humanos, proteção de dados e governança algorítmica. A delimitação temporal concentrou-se principalmente em trabalhos publicados entre 2016 e 2026, período caracterizado pela ampliação das aplicações de aprendizado de máquina e pela intensificação das discussões jurídicas sobre decisões automatizadas. Publicações anteriores foram consideradas

quando necessárias à compreensão dos fundamentos jurídicos internacionais.

A pesquisa bibliográfica foi complementada pela análise de documentos institucionais e normativos. Foram examinados instrumentos da Organização das Nações Unidas, UNESCO, Conselho da Europa, União Europeia, Organização dos Estados Americanos e Comissão Interamericana de Direitos Humanos. A análise compreendeu a Declaração Universal dos Direitos Humanos, os Pactos Internacionais de 1966, a Convenção Americana sobre Direitos Humanos, os Princípios Orientadores sobre Empresas e Direitos Humanos, a Recomendação da UNESCO sobre a Ética da Inteligência Artificial, relatórios do Alto Comissariado das Nações Unidas para os Direitos Humanos, a Resolução A/RES/78/265 da Assembleia Geral, a Convenção Quadro do Conselho da Europa sobre Inteligência Artificial e o Regulamento de Inteligência Artificial da União Europeia.

Os critérios de inclusão abrangeram trabalhos que discutissem diretamente os efeitos de sistemas algorítmicos sobre direitos reconhecidos internacionalmente, instrumentos institucionais que apresentassem princípios ou obrigações aplicáveis à Inteligência Artificial e documentos relacionados à responsabilização de Estados e empresas. Foram excluídos materiais predominantemente técnicos sem relação jurídica direta com o objeto, publicações de caráter exclusivamente comercial e textos que tratassem de inovação tecnológica sem estabelecer relação com Direitos Humanos ou mecanismos de proteção.

A análise foi organizada em categorias temáticas. A primeira reuniu os fundamentos internacionais aplicáveis à Inteligência Artificial. A

segunda concentrou os riscos relacionados à privacidade, vigilância e tratamento de dados. A terceira examinou igualdade, discriminação e efeitos de classificações automatizadas. A quarta abordou transparência, devido processo, supervisão humana e acesso à reparação. A quinta tratou da responsabilidade das empresas e dos problemas gerados pela atuação transnacional. A sexta comparou as respostas desenvolvidas pelos sistemas universal, europeu e interamericano.

Por se tratar de pesquisa bibliográfica e documental baseada em fontes públicas, não houve participação direta de seres humanos nem coleta de dados pessoais, razão pela qual não foram realizados procedimentos de intervenção ou experimentação. A principal limitação metodológica esteve associada à velocidade das transformações normativas e tecnológicas. Novos sistemas de Inteligência Artificial e novos instrumentos regulatórios podem modificar parte das discussões apresentadas. Essa condição, contudo, não comprometeu a análise dos fundamentos internacionais, pois o estudo concentrou-se em direitos e deveres jurídicos cuja aplicação permanece necessária independentemente da arquitetura tecnológica utilizada.

DESENVOLVIMENTO

Direitos Humanos como parâmetro jurídico para a Inteligência Artificial

A análise jurídica da Inteligência Artificial exige distinguir a tecnologia de suas aplicações concretas. A existência de um sistema automatizado não configura, isoladamente, uma ameaça aos Direitos Humanos. Os efeitos jurídicos dependem da finalidade para

a qual a tecnologia é utilizada, dos dados empregados, do nível de autonomia atribuído ao sistema, da possibilidade de supervisão e das consequências produzidas sobre indivíduos ou grupos. Essa distinção é importante porque impede respostas regulatórias baseadas apenas na novidade tecnológica e direciona a proteção para situações em que direitos efetivamente podem ser restringidos ou violados.

McGregor, Murray e Ng (2019) defendem que o Direito Internacional dos Direitos Humanos fornece uma estrutura útil para avaliar a utilização de algoritmos durante todo o seu ciclo de vida. Essa perspectiva modifica o foco da discussão. Em vez de analisar somente a precisão técnica de um sistema, torna-se necessário perguntar quais direitos podem ser afetados, qual agente possui obrigações, quais medidas de prevenção foram adotadas e quais mecanismos estão disponíveis para a pessoa atingida. A responsabilidade, portanto, não começa apenas após a ocorrência do dano.

Os direitos previstos internacionalmente foram formulados de maneira suficientemente ampla para acompanhar mudanças sociais e tecnológicas. A privacidade não protege apenas correspondências físicas, assim como a liberdade de expressão não permanece restrita a meios impressos. A interpretação dos tratados acompanha novas formas de interferência e novos ambientes de exercício dos direitos. Essa característica permite aplicar normas existentes à Inteligência Artificial sem depender da elaboração de um catálogo inteiramente novo de direitos digitais.

McGregor (2026) observa que os organismos vinculados ao sistema das Nações Unidas construíram um conjunto crescente de análises

sobre Inteligência Artificial e tecnologias relacionadas. A dificuldade passou a incluir a própria dispersão desse material entre diferentes órgãos, mandatos e documentos. A existência de múltiplas orientações não produz automaticamente um sistema coordenado de monitoramento. A proteção exige que as normas internacionais sejam incorporadas aos processos de desenvolvimento, contratação e utilização da tecnologia.

Kamble (2024) destaca que a Inteligência Artificial pode tanto favorecer quanto ameaçar Direitos Humanos, dependendo das condições de sua aplicação. Tecnologias empregadas para ampliar acesso a serviços de saúde, facilitar comunicação e melhorar acessibilidade podem contribuir para a efetivação de direitos. O mesmo conjunto de técnicas pode, porém, ser utilizado para vigilância excessiva, classificação discriminatória ou restrição indevida de oportunidades. A avaliação jurídica não deve assumir previamente uma posição favorável ou contrária à tecnologia, mas verificar seus efeitos e as salvaguardas existentes.

Essa abordagem também evita que princípios éticos sejam confundidos com obrigações jurídicas. Valores como transparência, equidade, responsabilidade e supervisão humana possuem relação direta com Direitos Humanos, mas a proteção internacional não pode depender apenas da adesão voluntária a códigos corporativos. Lane (2022) observa que iniciativas de governança contribuem para esclarecer padrões, porém as responsabilidades jurídicas continuam sendo essenciais quando tecnologias interferem em direitos protegidos por tratados.

A dignidade humana funciona como referência central nesse processo. Sistemas tecnológicos devem permanecer instrumentos a

serviço das pessoas e não mecanismos capazes de retirar completamente sua possibilidade de compreender, contestar ou influenciar decisões que produzem efeitos relevantes sobre sua vida. A proteção internacional torna-se especialmente necessária quando a capacidade tecnológica amplia assimetrias já existentes entre cidadãos, Estados e grandes organizações privadas.

Privacidade, vigilância e formação automatizada de perfis

A privacidade figura entre os direitos mais diretamente afetados pela expansão da Inteligência Artificial. O problema não está somente na quantidade de dados coletados, mas na capacidade de integrar informações provenientes de fontes distintas e utilizá-las para produzir inferências. Dados de localização, consumo, navegação, interação social ou comportamento podem adquirir novos significados quando submetidos a sistemas capazes de identificar padrões que seriam dificilmente percebidos por análise humana.

O Alto Comissariado das Nações Unidas para os Direitos Humanos (2021) identificou que sistemas de Inteligência Artificial podem ampliar interferências na privacidade por meio de reconhecimento facial, análise de comportamento, criação de perfis e tomada automatizada de decisões. A capacidade de processar grandes bases de dados permite produzir representações detalhadas da vida das pessoas. Essas informações podem afetar não apenas a privacidade, mas também liberdade de expressão, associação, acesso a serviços e igualdade.

Wachter e Mittelstadt (2019) chamam atenção para o problema das inferências automatizadas. O indivíduo pode fornecer determinados

dados para uma finalidade específica e, posteriormente, ser classificado a partir de conclusões produzidas pelo sistema. A informação derivada pode relacionar-se a risco financeiro, preferência política, saúde, comportamento profissional ou outras características. A dificuldade jurídica aparece quando essas inferências são tratadas como objetivas mesmo possuindo natureza probabilística.

A vigilância apoiada por Inteligência Artificial produz questão semelhante. Sistemas de reconhecimento facial e identificação biométrica ampliam a capacidade de localizar pessoas em espaços públicos e privados. A coleta de imagens isoladas já envolve questões relacionadas à privacidade, mas a possibilidade de identificação automatizada em grande escala transforma a dimensão do monitoramento. O poder deixa de depender da observação individual e passa a ser exercido por infraestruturas capazes de examinar milhares ou milhões de registros.

Matar e Murray (2025) relacionam a expansão da vigilância apoiada por Inteligência Artificial à liberdade necessária para construção da identidade individual. A percepção de estar constantemente submetido à observação pode modificar comportamentos mesmo quando nenhuma punição é efetivamente aplicada. A vigilância pode gerar efeitos inibidores sobre manifestação política, participação em movimentos sociais, escolha de grupos de convivência e exercício de outros direitos.

O princípio da proporcionalidade assume função relevante nesses casos. A existência de finalidade legítima não autoriza qualquer forma de tratamento automatizado. A medida deve ser adequada, necessária e proporcional à finalidade perseguida. O uso de

reconhecimento biométrico para localizar uma pessoa determinada em investigação específica apresenta características distintas da identificação permanente de todos os indivíduos que circulam por determinada área. A escala tecnológica modifica a intensidade da interferência.

A transparência também possui limites próprios. Informar genericamente que uma organização utiliza Inteligência Artificial não permite compreender quais dados foram empregados, que tipo de inferência foi produzida e como essa informação influenciou uma decisão. A proteção da privacidade exige medidas anteriores ao dano, incluindo limitação de finalidade, redução da coleta, segurança, controle de acesso e avaliação de riscos.

A questão se torna mais complexa quando empresas privadas acumulam grande quantidade de dados e disponibilizam serviços utilizados pelo próprio Estado. A separação tradicional entre vigilância pública e tratamento privado de informações pode tornar-se insuficiente quando organizações públicas contratam tecnologias desenvolvidas a partir de infraestruturas comerciais. O dever estatal de proteger direitos exige atenção à forma como esses contratos são estruturados e às garantias aplicáveis aos dados utilizados.

Igualdade, discriminação e vieses algorítmicos

A promessa de decisões tecnicamente neutras foi uma das justificativas para a adoção de sistemas automatizados em diferentes setores. Entretanto, a substituição parcial da decisão humana por modelos matemáticos não elimina necessariamente desigualdades. Algoritmos são desenvolvidos a partir de escolhas

humanas sobre dados, variáveis, objetivos, critérios de desempenho e resultados considerados aceitáveis.

Barocas e Selbst (2016) demonstram que bancos de dados podem carregar marcas de discriminações históricas. Quando um modelo aprende a partir desses registros, existe possibilidade de reproduzir padrões anteriores mesmo sem receber instruções discriminatórias explícitas. Em seleção profissional, por exemplo, dados históricos formados em ambiente no qual determinados grupos tiveram menor acesso a cargos podem levar o sistema a reconhecer como desejáveis características associadas ao grupo anteriormente predominante.

O problema é agravado pelo uso de variáveis substitutas. Mesmo quando raça, gênero ou outra característica protegida não é incluída no modelo, endereço, histórico escolar, padrões de consumo ou outras informações podem apresentar correlação com grupos sociais específicos. A exclusão formal de uma variável sensível não assegura neutralidade do resultado.

Mittelstadt et al. (2016) relacionam esse fenômeno a dificuldades de interpretação das decisões algorítmicas. Modelos complexos podem produzir classificações estatisticamente eficientes sem apresentar justificativa compreensível sobre a influência de cada variável. O desempenho agregado de um sistema também pode esconder diferenças relevantes entre grupos. Um modelo com elevado nível médio de precisão pode apresentar erros significativamente maiores para determinada população.

A discriminação algorítmica possui ainda dimensão coletiva. O Direito costuma oferecer instrumentos de contestação quando

determinada pessoa comprova tratamento indevido. Porém, modelos automatizados podem distribuir riscos de forma sistemática entre grupos, tornando insuficiente a análise de casos individuais. A proteção da igualdade exige examinar padrões de resultados e não apenas intenções declaradas pelos desenvolvedores.

Krupiy e Scheinin (2023) demonstram esse problema ao analisar a situação das pessoas com deficiência. Sistemas construídos com base em padrões considerados médios podem classificar como anormal uma característica associada à deficiência e transformar diferenças humanas em desvantagens automatizadas. A própria arquitetura da tecnologia pode gerar barreiras quando determinados modos de comunicação, mobilidade ou comportamento não estão adequadamente representados.

A Convenção sobre os Direitos das Pessoas com Deficiência oferece elementos relevantes porque exige inclusão, acessibilidade e igualdade de oportunidades. A interpretação desses direitos no ambiente digital demonstra que a Inteligência Artificial não pode ser avaliada apenas segundo sua eficiência para a maioria dos usuários. Um sistema que funciona adequadamente para grande parte da população pode continuar sendo incompatível com Direitos Humanos caso exclua de forma sistemática grupos específicos.

O dever de não discriminação exige medidas preventivas. Auditorias, testes com diferentes grupos, participação de comunidades potencialmente afetadas e avaliação de impacto podem identificar riscos antes da utilização em grande escala. A correção posterior permanece necessária quando o dano ocorre, mas não substitui o

dever de evitar que sistemas com problemas conhecidos sejam empregados em decisões relevantes.

A perspectiva internacional também é importante porque as bases de dados e os modelos podem ser desenvolvidos em contextos culturais diferentes daqueles em que serão utilizados. Um sistema treinado predominantemente com dados de determinadas regiões pode apresentar desempenho inferior quando aplicado em outros países. A transferência internacional de tecnologias sem avaliação das condições locais cria riscos específicos para populações pouco representadas durante o desenvolvimento.

Transparência, explicabilidade, devido processo e direito à contestação

A utilização de Inteligência Artificial em decisões relacionadas a direitos produz um problema que ultrapassa a precisão do resultado. Mesmo uma decisão estatisticamente adequada pode ser juridicamente questionável quando a pessoa afetada não consegue compreender os motivos que levaram à sua adoção. O devido processo exige possibilidade real de conhecimento e contestação.

McGregor, Murray e Ng (2019) consideram a transparência importante, mas insuficiente quando tratada de forma isolada. A disponibilização de informações técnicas não assegura responsabilização se o conteúdo apresentado for incompreensível para as pessoas atingidas. A transparência deve permitir algum nível de controle sobre o processo decisório, possibilitando identificar os critérios relevantes, verificar sua legitimidade e corrigir erros.

Essa distinção é particularmente importante em modelos complexos. A abertura integral de códigos e parâmetros nem

sempre produz explicação útil. Em determinadas situações, o interessado necessita saber quais informações influenciaram a decisão e como pode demonstrar que estão incorretas. O direito relevante não é necessariamente compreender matematicamente toda a estrutura computacional, mas receber justificativa suficiente para exercer defesa efetiva.

Rachovitsa e Johann (2022) demonstram que a opacidade pode comprometer também o controle judicial. Se órgãos públicos utilizam sistemas cujos critérios permanecem inacessíveis, o próprio Poder Judiciário encontra dificuldades para avaliar legalidade e proporcionalidade. A proteção processual exige que autoridades públicas mantenham capacidade de explicar as decisões tomadas com apoio tecnológico.

A supervisão humana costuma ser apresentada como solução, mas sua eficácia depende das condições concretas de funcionamento. A existência formal de um servidor ou funcionário responsável por confirmar uma recomendação automatizada não garante análise independente. Quando o profissional possui pouco tempo, desconhece as limitações do sistema ou tende a confiar automaticamente no resultado apresentado, a revisão humana pode funcionar apenas como validação formal.

A supervisão precisa ser significativa. A pessoa responsável deve possuir autoridade para alterar a decisão, acesso às informações necessárias e conhecimento suficiente para identificar situações nas quais o modelo pode falhar. Também deve existir registro do processo para permitir controle posterior.

O direito à contestação adquire importância especial em decisões relacionadas a benefícios sociais, emprego, crédito, saúde, educação, imigração e segurança. Nessas áreas, classificações incorretas podem produzir consequências que não são facilmente reversíveis. Uma pessoa excluída indevidamente de determinada oportunidade pode sofrer efeitos econômicos e sociais mesmo que a decisão seja corrigida posteriormente.

A reparação efetiva constitui elemento central do sistema internacional de proteção. Não basta reconhecer abstratamente que determinada tecnologia causou dano. Devem existir procedimentos capazes de interromper a violação, corrigir registros, rever decisões e, quando apropriado, compensar a vítima. Essa exigência enfrenta dificuldades particulares quando vários agentes participaram do funcionamento do sistema.

Responsabilidade das empresas e fragmentação da cadeia tecnológica

Grande parte da infraestrutura de Inteligência Artificial é desenvolvida, controlada ou operada por empresas privadas. Esse elemento diferencia parte dos riscos tecnológicos de violações tradicionalmente associadas à atuação direta do Estado. Plataformas digitais, fornecedores de computação, empresas de dados e desenvolvedores de modelos participam de cadeias que atravessam diferentes países e setores econômicos.

Os Princípios Orientadores sobre Empresas e Direitos Humanos estabeleceram uma estrutura baseada no dever estatal de proteger, na responsabilidade empresarial de respeitar e na necessidade de acesso a mecanismos de reparação. Embora não transformem as

empresas em partes dos principais tratados de Direitos Humanos, os Princípios contribuíram para consolidar a expectativa de que organizações privadas realizem diligência adequada para identificar, prevenir e reduzir impactos negativos.

O Alto Comissariado das Nações Unidas para os Direitos Humanos (2025) aplicou essa lógica de maneira específica às empresas de tecnologia e às atividades relacionadas à Inteligência Artificial. A diligência em Direitos Humanos deve acompanhar o ciclo de vida dos produtos e considerar não apenas a atuação direta da empresa, mas também relações comerciais e modelos de negócio capazes de ampliar riscos.

O Grupo de Trabalho das Nações Unidas sobre Empresas e Direitos Humanos (2025) destacou que o problema não se limita às empresas que desenvolvem modelos. Estados e organizações privadas que adquirem e implantam sistemas criados por terceiros também possuem responsabilidades. A contratação de tecnologia não permite transferir integralmente ao fornecedor a obrigação de avaliar consequências sobre Direitos Humanos.

Esse aspecto possui consequência direta para contratos públicos. Um órgão estatal que utiliza sistema de Inteligência Artificial para apoiar decisões continua responsável por assegurar legalidade, igualdade, privacidade e possibilidade de contestação. Alegar desconhecimento sobre o funcionamento de tecnologia contratada não elimina o dever público de proteção.

A distribuição internacional das atividades empresariais dificulta a obtenção de reparação. Desenvolvimento, armazenamento, tratamento de dados e utilização podem ocorrer em jurisdições

distintas. Pessoas atingidas nem sempre conseguem identificar qual empresa participou de determinada etapa ou qual legislação deve ser aplicada. Os mecanismos internacionais de proteção enfrentam, assim, uma estrutura empresarial global que não corresponde integralmente às fronteiras tradicionais da jurisdição estatal.

Lane (2022) aponta que a responsabilidade das empresas continua sendo uma área de incerteza relevante no Direito Internacional dos Direitos Humanos aplicado à Inteligência Artificial. A multiplicidade de iniciativas éticas ajuda a construir padrões, mas não resolve integralmente questões relativas à imposição de deveres, investigação de violações e execução de decisões contra atores transnacionais.

A resposta exige combinação de instrumentos. Estados devem regulamentar atividades empresariais sob sua jurisdição, estabelecer obrigações de diligência, exigir avaliações de impacto e assegurar mecanismos judiciais e administrativos acessíveis. Sistemas internacionais podem contribuir por meio da interpretação dos tratados, monitoramento, recomendações e desenvolvimento de padrões comuns capazes de reduzir lacunas entre diferentes jurisdições.

O sistema universal das Nações Unidas e a construção de parâmetros internacionais

O sistema universal possui papel relevante porque reúne tratados, órgãos de monitoramento, procedimentos especiais e instituições capazes de interpretar os Direitos Humanos em escala global. Sua principal vantagem está na possibilidade de estabelecer referências que ultrapassem modelos regulatórios nacionais ou regionais.

A Declaração Universal dos Direitos Humanos e os Pactos Internacionais de 1966 fornecem bases para examinar diferentes impactos da Inteligência Artificial. Privacidade, igualdade, liberdade de expressão, liberdade de associação, trabalho, educação, participação pública e acesso à justiça podem ser afetados por tecnologias automatizadas. O fato de esses instrumentos não mencionarem algoritmos não reduz sua aplicação.

O Alto Comissariado das Nações Unidas para os Direitos Humanos (2021) contribuiu para tornar mais concreta essa relação ao examinar Inteligência Artificial e privacidade. O relatório destacou que a tecnologia pode ampliar interferências em direitos e defendeu salvaguardas proporcionais aos riscos. A abordagem possui importância porque afasta a ideia de que a governança tecnológica deve ser tratada apenas como questão de inovação ou segurança cibernética.

A UNESCO (2021) ampliou esse debate por meio de sua Recomendação sobre a Ética da Inteligência Artificial. O documento combinou valores gerais com áreas de ação relacionadas a dados, educação, saúde, ambiente, gênero, comunicação e informação. Seu alcance político é relevante porque apresenta parâmetros comuns para Estados com diferentes níveis de desenvolvimento tecnológico.

A Assembleia Geral das Nações Unidas (2024) reforçou a necessidade de sistemas seguros e confiáveis e chamou atenção para as desigualdades existentes entre países. Esse ponto possui importância particular para os Direitos Humanos. A governança internacional da Inteligência Artificial não pode ser construída somente a partir das condições econômicas e tecnológicas dos

países que concentram infraestrutura, empresas e centros de pesquisa.

As diferenças de capacidade regulatória podem produzir assimetrias. Países dependentes de tecnologias desenvolvidas no exterior possuem menor influência sobre padrões técnicos e modelos empresariais. A proteção internacional necessita considerar essa distribuição desigual de recursos e conhecimento para evitar que determinados Estados permaneçam apenas como consumidores de sistemas cujos riscos não conseguem avaliar adequadamente.

Os documentos das Nações Unidas publicados em 2025 sobre empresas e Inteligência Artificial demonstraram uma transição importante da formulação de princípios para a discussão de mecanismos de diligência. O Alto Comissariado das Nações Unidas para os Direitos Humanos (2025) enfatizou que empresas tecnológicas devem incorporar Direitos Humanos às decisões relacionadas a modelos de negócio, projeto de produtos e relações comerciais. O Grupo de Trabalho das Nações Unidas sobre Empresas e Direitos Humanos (2025) estendeu a análise à aquisição e implantação de sistemas por Estados e empresas que não são desenvolvedores.

Apesar desses avanços, o sistema universal enfrenta limitações. Grande parte dos instrumentos específicos sobre Inteligência Artificial possui natureza recomendatória. Mesmo tratados obrigatórios dependem de implementação doméstica, cooperação estatal e mecanismos de monitoramento. A ausência de uma jurisdição universal com competência geral para julgar violações individuais reduz a capacidade de produzir respostas uniformes.

A multiplicidade de órgãos também pode gerar dispersão. McGregor (2026) observa que o sistema das Nações Unidas desenvolveu considerável produção sobre tecnologias digitais, mas ainda precisa aproveitar melhor sua capacidade de monitoramento. A questão deixou de ser apenas formular princípios e passou a exigir acompanhamento da forma como Estados e empresas incorporam esses parâmetros.

O sistema europeu e a formação de obrigações específicas

O sistema europeu apresenta atualmente a estrutura internacional mais desenvolvida em relação à Inteligência Artificial. Essa condição resulta da interação entre a Convenção Europeia de Direitos Humanos, a jurisprudência regional, normas de proteção de dados, legislação da União Europeia e a Convenção Quadro do Conselho da Europa.

A Convenção Europeia de Direitos Humanos oferece garantias relacionadas à vida privada, liberdade de expressão, igualdade e devido processo. A aplicação evolutiva desses direitos permite alcançar novas formas de interferência tecnológica. Casos relacionados à vigilância e ao tratamento de informações construíram fundamentos que podem ser utilizados na avaliação da Inteligência Artificial.

A experiência analisada por Rachovitsa e Johann (2022) demonstra a relevância dessa estrutura. No caso SyRI, o uso de sistema governamental destinado à identificação de possíveis fraudes em benefícios sociais foi confrontado com exigências relativas à vida privada e proporcionalidade. O episódio evidenciou que políticas

públicas baseadas em análise automatizada continuam submetidas às garantias de Direitos Humanos.

A Convenção Quadro do Conselho da Europa sobre Inteligência Artificial, Direitos Humanos, Democracia e Estado de Direito representou mudança qualitativa. O Conselho da Europa (2024) adotou instrumento vinculante especificamente voltado ao ciclo de vida da Inteligência Artificial. A Convenção estabelece princípios de dignidade humana, autonomia, transparência, supervisão, responsabilidade, igualdade, privacidade e confiabilidade.

Outro elemento relevante é a previsão de avaliações de risco e impacto. Essa abordagem procura identificar problemas antes que o dano esteja consolidado. A lógica preventiva aproxima a governança tecnológica da própria estrutura dos Direitos Humanos, pois Estados possuem não apenas dever de reparar violações, mas obrigações positivas de proteção.

O Regulamento de Inteligência Artificial da União Europeia adotou modelo baseado em níveis de risco. A União Europeia (2024) estabeleceu categorias de utilização e requisitos diferenciados, incluindo proibições para algumas práticas e obrigações específicas para sistemas de alto risco. A norma não substitui os tratados de Direitos Humanos, mas cria instrumentos técnicos e administrativos capazes de contribuir para sua implementação.

Existe, contudo, diferença institucional entre Conselho da Europa e União Europeia. O primeiro possui estrutura própria de Direitos Humanos e alcance geográfico distinto. A União Europeia, por sua vez, possui competências regulatórias e econômicas mais amplas sobre seus Estados membros. A coexistência desses instrumentos

pode fortalecer a proteção, embora também exija coordenação para evitar interpretações divergentes.

A experiência europeia oferece referências importantes, mas não deve ser simplesmente reproduzida em outras regiões. Contextos econômicos, institucionais e tecnológicos distintos exigem adaptações. O valor do modelo está principalmente na demonstração de que princípios de Direitos Humanos podem ser convertidos em obrigações relacionadas a avaliação de riscos, documentação, supervisão e possibilidade de recurso.

O sistema interamericano diante dos impactos da Inteligência Artificial

O sistema interamericano possui fundamentos normativos capazes de alcançar situações relacionadas à Inteligência Artificial, ainda que não exista, no âmbito da Organização dos Estados Americanos, instrumento equivalente à Convenção Quadro do Conselho da Europa. A Convenção Americana sobre Direitos Humanos protege garantias judiciais, honra, dignidade, liberdade de consciência, liberdade de expressão, igualdade e proteção judicial, direitos que podem ser diretamente afetados pela utilização de sistemas automatizados.

A região possui características próprias que tornam essa discussão particularmente relevante. Desigualdade social, diferenças de acesso à tecnologia, fragilidade institucional em determinados contextos e concentração econômica podem aumentar os efeitos negativos de decisões automatizadas. Sistemas desenvolvidos em outros ambientes sociais também podem não representar adequadamente a diversidade da população latino-americana.

A Comissão Interamericana de Direitos Humanos e sua Relatoria Especial para a Liberdade de Expressão passaram a acompanhar com maior atenção o ambiente digital. A Comissão Interamericana de Direitos Humanos (2025) registrou preocupações relacionadas ao avanço da Inteligência Artificial generativa, à manipulação informacional, aos vieses, à discriminação, à privacidade e à proteção de dados. A presença dessas questões nos relatórios regionais demonstra que os efeitos da IA já alcançaram o trabalho de monitoramento do sistema.

A liberdade de expressão constitui campo especialmente relevante. Sistemas de recomendação definem parte significativa das informações apresentadas aos usuários de plataformas digitais. Moderação automatizada pode reduzir alcance, remover conteúdos ou classificar publicações. Ao mesmo tempo, ferramentas de Inteligência Artificial generativa ampliam a capacidade de produzir conteúdos sintéticos capazes de afetar reputação, debate público e acesso à informação confiável.

A proteção interamericana precisa evitar duas respostas extremas. A primeira seria considerar as plataformas e seus sistemas algorítmicos inteiramente alheios às obrigações de Direitos Humanos. A segunda seria utilizar a proteção de direitos como justificativa genérica para controles estatais desproporcionais sobre circulação de informações. A tradição interamericana de proteção da liberdade de expressão exige critérios de legalidade, finalidade legítima, necessidade e proporcionalidade.

A igualdade também demanda atenção. O uso de Inteligência Artificial em segurança pública, políticas sociais e seleção de beneficiários pode atingir grupos que historicamente enfrentam

discriminação. Caso modelos sejam construídos com dados decorrentes de práticas institucionais desiguais, a automação poderá conferir aparência técnica à continuidade desses padrões.

O sistema interamericano dispõe de instrumentos que podem ser desenvolvidos por interpretação. Relatórios temáticos, medidas cautelares, análise de petições e jurisprudência da Corte Interamericana podem gradualmente delimitar deveres relacionados às decisões automatizadas. A experiência construída em matérias como privacidade, igualdade, devido processo e liberdade de expressão oferece fundamentos para essa evolução.

A ausência de instrumento regional específico, portanto, não significa ausência de proteção. O principal desafio é produzir parâmetros suficientemente claros para orientar os Estados da região. Avaliação de impacto em Direitos Humanos, transparência compatível com o contexto, mecanismos de contestação e proteção de grupos vulneráveis constituem áreas nas quais o sistema interamericano pode desenvolver padrões próprios.

Cooperação internacional, desigualdade tecnológica e acesso à reparação

A dimensão internacional da Inteligência Artificial impede que a proteção seja resolvida apenas por legislação doméstica. Sistemas desenvolvidos em um país podem ser utilizados simultaneamente em dezenas de outros. Dados atravessam fronteiras e serviços digitais são oferecidos sem presença física significativa em todos os territórios alcançados.

A fragmentação regulatória pode produzir lacunas. Empresas podem enfrentar obrigações distintas conforme a jurisdição,

enquanto usuários possuem diferentes níveis de proteção. A harmonização completa não é necessariamente possível ou desejável, mas padrões internacionais mínimos podem impedir que diferenças regulatórias sejam utilizadas para reduzir garantias.

A cooperação precisa incluir intercâmbio técnico entre autoridades, compatibilidade entre avaliações de impacto e mecanismos de assistência destinados aos países com menor capacidade institucional. A simples imposição de obrigações complexas sem estrutura para sua implementação pode aumentar desigualdades entre Estados.

A desigualdade tecnológica também influencia a participação na construção das normas. Países que concentram empresas, infraestrutura computacional e pesquisa possuem maior capacidade de definir padrões. A governança internacional precisa assegurar que países em desenvolvimento participem das decisões que influenciarão tecnologias utilizadas em seus próprios territórios.

Outro desafio está relacionado à reparação. A pessoa prejudicada por uma decisão algorítmica precisa saber que foi submetida a esse processo, identificar o agente responsável e encontrar mecanismo acessível para contestá-lo. Nenhuma dessas etapas pode ser presumida. Sistemas opacos dificultam a percepção do próprio dano e cadeias empresariais complexas dificultam a identificação do responsável.

A proteção efetiva depende, portanto, de direitos processuais. Aviso sobre utilização de decisões automatizadas, acesso a informações relevantes, revisão por pessoa qualificada, possibilidade de apresentar provas, correção de dados e acesso a órgãos

independentes são mecanismos que aproximam os princípios internacionais da experiência concreta dos indivíduos.

O desafio central dos sistemas internacionais consiste em transformar princípios amplos em garantias operacionais. A afirmação de que Inteligência Artificial deve respeitar Direitos Humanos possui pouca utilidade se não for acompanhada de procedimentos capazes de verificar seu cumprimento. Avaliação de impacto, documentação, auditorias, supervisão independente e reparação devem integrar a estrutura jurídica da tecnologia.

RESULTADOS E DISCUSSÃO

A análise das fontes permitiu identificar que o primeiro resultado relevante consiste na existência de uma base normativa internacional já aplicável aos principais riscos decorrentes da Inteligência Artificial. Direitos relacionados à privacidade, igualdade, liberdade de expressão, devido processo e proteção judicial não dependem de reconhecimento como novos direitos digitais para incidirem sobre decisões automatizadas. A principal dificuldade está na adaptação dos mecanismos de implementação às características dos sistemas tecnológicos. Esse resultado converge com McGregor, Murray e Ng (2019), para quem o Direito Internacional dos Direitos Humanos oferece estrutura útil para responsabilização algorítmica, embora ainda existam dificuldades relacionadas à atuação empresarial e à aplicação transnacional das obrigações.

O segundo resultado refere-se à insuficiência de respostas baseadas exclusivamente em transparência. A literatura analisada demonstrou que acesso a informações técnicas não garante capacidade de compreender e contestar decisões. Mittelstadt et al. (2016)

identificam que complexidade e autonomia algorítmica dificultam a atribuição de responsabilidade, enquanto Wachter e Mittelstadt (2019) mostram que inferências automatizadas introduzem problemas que ultrapassam a simples coleta de dados. A proteção demanda transparência orientada ao exercício de direitos, associada a explicação adequada, supervisão e possibilidade de revisão.

O terceiro resultado revela que igualdade e não discriminação ocupam posição central na discussão. Barocas e Selbst (2016) demonstraram que dados históricos podem reproduzir desigualdades preexistentes, e Krupiy e Scheinin (2023) evidenciaram que sistemas automatizados podem criar barreiras específicas para pessoas com deficiência. Isso significa que a precisão média de uma tecnologia não constitui critério suficiente para determinar sua compatibilidade com Direitos Humanos. A avaliação precisa considerar distribuição de erros, impactos sobre grupos e condições de acessibilidade.

O quarto resultado foi a existência de assimetria entre os sistemas internacionais analisados. O sistema universal produziu quantidade expressiva de recomendações, relatórios e orientações, mas permanece fortemente dependente da implementação estatal. O sistema europeu avançou para um instrumento vinculante específico, a Convenção Quadro de 2024, além de possuir regulamentação detalhada no âmbito da União Europeia. O sistema interamericano possui direitos e mecanismos capazes de alcançar violações relacionadas à IA, mas ainda demanda maior sistematização de parâmetros específicos para decisões automatizadas.

A análise também identificou que a responsabilidade empresarial representa uma das áreas mais complexas. O Alto Comissariado das Nações Unidas para os Direitos Humanos (2025) e o Grupo de Trabalho das Nações Unidas sobre Empresas e Direitos Humanos (2025) reforçaram a necessidade de diligência ao longo do ciclo de vida da tecnologia. O avanço é relevante porque amplia a análise para além do desenvolvedor e inclui organizações que comprem, integram e utilizam sistemas de terceiros.

Os resultados indicaram ainda que a existência de regulação específica não elimina a importância dos tratados tradicionais de Direitos Humanos. Ao contrário, os instrumentos mais recentes tendem a utilizar como referência direitos já consolidados. O Conselho da Europa (2024), por exemplo, estruturou sua Convenção Quadro sobre dignidade, autonomia, igualdade, privacidade, responsabilidade e Estado de Direito. A inovação regulatória ocorre principalmente nos mecanismos de operacionalização desses direitos.

Entre as limitações identificadas, permanece a velocidade da transformação tecnológica. Modelos e aplicações podem modificar-se em período menor que o necessário à negociação de tratados internacionais. Essa diferença temporal favorece instrumentos baseados em direitos e riscos, capazes de permanecer aplicáveis mesmo quando determinada tecnologia é substituída. A regulação excessivamente vinculada a uma arquitetura específica tende a perder eficácia mais rapidamente.

A discussão demonstra, por fim, que o problema fundamental não consiste em escolher entre proteção de Direitos Humanos e inovação tecnológica. A incorporação de avaliações de impacto,

mecanismos de contestação e responsabilidades definidas pode contribuir para sistemas mais confiáveis. A inovação juridicamente sustentável depende da capacidade de reconhecer antecipadamente os riscos e impedir que ganhos de eficiência sejam obtidos por meio da redução de garantias fundamentais.

CONSIDERAÇÕES FINAIS

O estudo analisou os limites e os desafios enfrentados pelos sistemas internacionais de proteção dos Direitos Humanos diante da expansão da Inteligência Artificial. A investigação demonstrou que a transformação tecnológica não eliminou a aplicação das normas internacionais existentes. Privacidade, igualdade, liberdade de expressão, devido processo, dignidade e direito à reparação permanecem juridicamente aplicáveis quando decisões são tomadas ou apoiadas por sistemas automatizados.

A questão inicialmente formulada buscou compreender em que medida os sistemas internacionais possuem instrumentos capazes de responder às violações e aos riscos associados à Inteligência Artificial. Os resultados permitem afirmar que existe base normativa suficiente para reconhecer grande parte desses riscos como problemas de Direitos Humanos. Entretanto, essa capacidade normativa não se converte automaticamente em proteção efetiva. O principal desafio está na aplicação dos direitos às características técnicas, empresariais e transnacionais dos sistemas atuais.

A opacidade algorítmica constitui uma das dificuldades mais significativas. Pessoas afetadas por sistemas automatizados podem desconhecer que determinado resultado decorreu da utilização de Inteligência Artificial. Mesmo quando essa informação é

disponibilizada, pode permanecer incerto quais dados foram considerados, como foram interpretados e por que determinada classificação foi produzida. Sem acesso mínimo a esses elementos, o direito de contestar uma decisão pode tornar-se meramente formal.

A transparência, contudo, não deve ser tratada como solução isolada. A divulgação de código fonte, parâmetros ou documentação técnica pode ser útil para especialistas e autoridades de fiscalização, mas a pessoa afetada necessita de informação relacionada à decisão concreta. O desafio jurídico consiste em estabelecer níveis de explicação adequados a cada finalidade, preservando direitos sem exigir que todos os destinatários possuam conhecimento técnico especializado.

A igualdade apresenta problema distinto. Algoritmos podem reproduzir desigualdades sem utilizar explicitamente características protegidas. Dados históricos, variáveis correlacionadas e diferenças de representação entre grupos são suficientes para produzir impactos desproporcionais. A ausência de intenção discriminatória do desenvolvedor não elimina o efeito sobre a pessoa atingida. Os sistemas internacionais precisam, por isso, fortalecer uma compreensão de discriminação capaz de examinar resultados e estruturas, e não apenas intenções.

O estudo também constatou que grupos vulneráveis exigem atenção específica. Pessoas com deficiência, mulheres, minorias raciais, migrantes, populações de baixa renda e outros grupos sujeitos a desigualdades históricas podem sofrer efeitos mais intensos quando modelos são desenvolvidos a partir de padrões que não representam adequadamente suas experiências. A participação desses grupos nas avaliações de impacto deve ser considerada

componente da prevenção e não procedimento meramente consultivo.

A responsabilização empresarial permanece um dos pontos mais delicados. Empresas possuem papel central no desenvolvimento de modelos, fornecimento de infraestrutura, tratamento de dados e oferta de serviços. Ao mesmo tempo, o Direito Internacional dos Direitos Humanos foi historicamente estruturado com foco nas obrigações estatais. Os Princípios Orientadores sobre Empresas e Direitos Humanos forneceram caminho importante ao consolidar responsabilidade empresarial de respeito e diligência, mas persistem dificuldades de exigibilidade e reparação em cadeias transnacionais.

Os documentos das Nações Unidas publicados em 2025 demonstram avanço nesse ponto ao relacionar expressamente a diligência em Direitos Humanos às atividades de empresas de tecnologia e à contratação de sistemas de Inteligência Artificial. A responsabilidade não pode ser limitada ao desenvolvedor original. Organizações que escolhem implantar uma tecnologia e utilizam seus resultados em decisões também devem verificar sua adequação, limitações e riscos.

A experiência do Conselho da Europa constitui passo importante para superar parte da fragmentação normativa. A Convenção Quadro adotada em 2024 demonstrou que é juridicamente possível construir instrumento internacional vinculante específico para Inteligência Artificial sem abandonar a centralidade dos direitos já reconhecidos. O modelo baseado em dignidade, igualdade, privacidade, transparência, responsabilidade e avaliação de riscos oferece referência para outras regiões.

A existência da Convenção não significa, entretanto, que os desafios estejam resolvidos. A eficácia de um tratado depende de adesão, incorporação e implementação pelos Estados. Além disso, a velocidade de desenvolvimento da Inteligência Artificial exige mecanismos capazes de acompanhar mudanças técnicas sem necessidade de renegociação permanente do texto convencional. A escolha por princípios e avaliações de risco procura responder a essa dificuldade.

No sistema universal, a produção normativa e institucional apresentou crescimento considerável. A UNESCO, o Alto Comissariado das Nações Unidas para os Direitos Humanos, a Assembleia Geral e outros órgãos desenvolveram orientações destinadas a aproximar a governança de IA dos Direitos Humanos. A etapa seguinte exige maior coordenação entre esses documentos e os mecanismos tradicionais de monitoramento.

A situação do sistema interamericano apresenta oportunidade de desenvolvimento. A Convenção Americana contém fundamentos suficientes para enfrentar problemas relacionados à IA, e a Comissão Interamericana já passou a observar os impactos das tecnologias digitais sobre liberdade de expressão, privacidade, igualdade e circulação de informações. Ainda assim, seria importante consolidar parâmetros regionais específicos relacionados a avaliações de impacto, decisões automatizadas, reconhecimento biométrico, supervisão humana e direito de contestação.

Essa construção precisa considerar as particularidades econômicas e sociais da América Latina. Soluções desenvolvidas para países com infraestrutura tecnológica e institucional distinta não podem ser simplesmente transplantadas. A região necessita de parâmetros que

preservem Direitos Humanos sem impedir usos socialmente benéficos da tecnologia e sem criar requisitos que apenas grandes organizações possam cumprir.

Outro resultado importante consiste na necessidade de cooperação internacional. A circulação transfronteiriça de dados e serviços torna insuficiente uma proteção limitada às fronteiras nacionais. Empresas podem operar simultaneamente em diferentes jurisdições, enquanto modelos desenvolvidos em determinado país são utilizados em contextos sociais muito distintos. Padrões mínimos comuns podem reduzir espaços de desproteção.

A avaliação de impacto em Direitos Humanos aparece como instrumento especialmente adequado. Realizada antes da implantação e atualizada durante a utilização, ela permite identificar pessoas e grupos afetados, avaliar a gravidade dos riscos e estabelecer medidas de prevenção. Sua utilidade depende, contudo, de não ser reduzida a procedimento burocrático. Avaliações precisam produzir consequências reais, inclusive a decisão de não utilizar determinado sistema quando seus riscos não puderem ser adequadamente controlados.

O mesmo raciocínio se aplica à supervisão humana. A mera presença de uma pessoa no fluxo decisório não assegura proteção. Para ser efetiva, a supervisão exige conhecimento, autonomia e condições para contrariar o resultado algorítmico. A autoridade humana não pode funcionar apenas como confirmação automática de uma classificação produzida pelo sistema.

A reparação deve ocupar posição equivalente à prevenção. Quando um sistema produz violação, é necessário assegurar revisão da

decisão, interrupção da prática, correção de dados e reparação adequada. A multiplicidade de empresas e instituições envolvidas não pode resultar em ausência de responsável. A arquitetura jurídica deve acompanhar a arquitetura tecnológica e distribuir deveres conforme o controle exercido por cada agente.

Concluiu se, portanto, que o desafio colocado pela Inteligência Artificial aos sistemas internacionais de proteção não consiste principalmente na necessidade de criar uma nova geração de Direitos Humanos. O problema está em garantir que direitos já reconhecidos permaneçam efetivos quando o exercício do poder passa a ocorrer por meio de dados, modelos estatísticos e decisões automatizadas.

A proteção internacional deverá combinar normas tradicionais, instrumentos específicos, avaliação preventiva, responsabilidade estatal, diligência empresarial e mecanismos acessíveis de reparação. A inovação pode contribuir para saúde, educação, acessibilidade, serviços públicos e produção científica, mas seus benefícios não justificam a redução de garantias fundamentais. A tecnologia deve permanecer submetida ao Direito e à dignidade humana, e não o contrário.

O fortalecimento dos sistemas internacionais dependerá, por fim, da capacidade de transformar princípios em obrigações concretas. Transparência, equidade, supervisão e responsabilidade somente produzirão proteção quando puderem ser verificadas, exigidas e utilizadas pelas pessoas afetadas. A Inteligência Artificial modifica os instrumentos pelos quais decisões são produzidas, mas não altera a finalidade central dos Direitos Humanos: limitar o exercício arbitrário

do poder e assegurar condições para que cada pessoa possa exercer seus direitos com liberdade, igualdade e dignidade.

REFERÊNCIAS BIBLIOGRÁFICAS

BAROCAS, Solon; SELBST, Andrew D. Big Data's disparate impact. **California Law Review**, Berkeley, v. 104, n. 3, p. 671-732, 2016. DOI: 10.15779/Z38BG31.

COUNCIL OF EUROPE. **Convention for the Protection of Human Rights and Fundamental Freedoms**. Rome: Council of Europe, 1950.

COUNCIL OF EUROPE. **Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law**. CETS No. 225. Strasbourg: Council of Europe, 2024.

EUROPEAN UNION. **Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence**. Official Journal of the European Union, 12 jul. 2024.

INTER-AMERICAN COMMISSION ON HUMAN RIGHTS. **Annual Report of the Office of the Special Rapporteur for Freedom of Expression, 2024**. OEA/Ser.L/V/II, Doc. 39/25, v. 2. Washington, D.C.: Organization of American States, 2025.

KAMBLE, R. M. Artificial intelligence and human rights. **Uniform Law Review**, Oxford, v. 29, n. 1, p. 77-86, 2024. DOI: 10.1093/ulr/unae020.

KRUPIY, Tetyana; SCHEININ, Martin. Disability discrimination in the digital realm: how the ICRPD applies to artificial intelligence decision making processes and helps in determining the state of

international human rights law. **Human Rights Law Review**, Oxford, v. 23, n. 3, ngad019, 2023. DOI: 10.1093/hrlr/ngad019.

LANE, Lottie. Clarifying human rights standards through artificial intelligence initiatives. **International & Comparative Law Quarterly**, Cambridge, v. 71, n. 4, p. 915-944, 2022. DOI: 10.1017/S0020589322000380.

MATAR, Ronit; MURRAY, Daragh. Re thinking international human rights law's approach to identity in light of surveillance and AI. **Human Rights Law Review**, Oxford, v. 25, n. 3, ngaf016, 2025. DOI: 10.1093/hrlr/ngaf016.

MCGREGOR, Lorna; MURRAY, Daragh; NG, Vivian. International human rights law as a framework for algorithmic accountability. **International & Comparative Law Quarterly**, Cambridge, v. 68, n. 2, p. 309-343, 2019. DOI: 10.1017/S0020589319000046.

MCGREGOR, Lorna. Artificial intelligence and human rights. In: GENSER, Jared; AL HUSSEIN, Zeid Ra'ad (ed.). **The Oxford Handbook of the UN Human Rights System**. Oxford: Oxford University Press, 2026. p. 1188-1218. DOI: 10.1093/9780197599983.003.0042.

MITTELSTADT, Brent Daniel; ALLO, Patrick; TADDEO, Mariarosaria; WACHTER, Sandra; FLORIDI, Luciano. The ethics of algorithms: mapping the debate. **Big Data & Society**, London, v. 3, n. 2, p. 1-21, 2016. DOI: 10.1177/2053951716679679.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Alto Comissariado das Nações Unidas para os Direitos Humanos. **The right to privacy in the digital age**. A/HRC/48/31. Geneva: United Nations, 2021.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Assembleia Geral. **Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development.** Resolution A/RES/78/265. New York: United Nations, 2024.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Declaração Universal dos Direitos Humanos.** Paris: ONU, 1948.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Guiding Principles on Business and Human Rights: implementing the United Nations “Protect, Respect and Remedy” Framework.** New York; Geneva: United Nations, 2011.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Human Rights Council. Working Group on the Issue of Human Rights and Transnational Corporations and Other Business Enterprises. **Artificial intelligence procurement and deployment: ensuring alignment with the Guiding Principles on Business and Human Rights.** A/HRC/59/53. Geneva: United Nations, 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. Office of the High Commissioner for Human Rights. **Practical application of the Guiding Principles on Business and Human Rights to the activities of technology companies, including activities relating to artificial intelligence.** A/HRC/59/32. Geneva: United Nations, 2025.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Pacto Internacional sobre Direitos Civis e Políticos.** Nova York: ONU, 1966.

ORGANIZAÇÃO DAS NAÇÕES UNIDAS. **Pacto Internacional sobre Direitos Econômicos, Sociais e Culturais.** Nova York: ONU, 1966.

ORGANIZAÇÃO DOS ESTADOS AMERICANOS. **Convenção Americana sobre Direitos Humanos.** San José: OEA, 1969.

RACHOVITSA, Adamantia; JOHANN, Niclas. The human rights implications of the use of AI in the digital welfare state: lessons learned from the Dutch SyRI case. **Human Rights Law Review**, Oxford, v. 22, n. 2, ngac010, 2022. DOI: 10.1093/hrlr/ngac010.

UNESCO. **Recommendation on the Ethics of Artificial Intelligence.** Paris: United Nations Educational, Scientific and Cultural Organization, 2021.

WACHTER, Sandra; MITTELSTADT, Brent. A right to reasonable inferences: re thinking data protection law in the age of Big Data and AI. **Columbia Business Law Review**, New York, v. 2019, n. 2, p. 494-620, 2019. DOI: 10.7916/cblr.v2019i2.3424.

¹ Graduação em Direito pelo Centro Universitário UniJorge (Salvador/Brasil). Pós-Graduada em Direito Público pelo Instituto Excelência JUSPODIVM (Salvador/Brasil). Mestranda em Estudos Jurídicos com ênfase em Direito Internacional - MUST University (MUST)