

DIREITO E INTELIGÊNCIA ARTIFICIAL NA (NÃO) REDUÇÃO DE DESIGUALDADES GLOBAIS: DECISÕES AUTOMATIZADAS NA IMIGRAÇÃO E SISTEMAS DE REFUGIADOS

*LAW AND ARTIFICIAL INTELLIGENCE IN THE (NON) REDUCTION OF
GLOBAL INEQUALITIES: AUTOMATED DECISIONS IN IMMIGRATION
AND REFUGEE SYSTEMS*

Fabiano Hartmann Peixoto¹ 

¹Universidade de Brasília, Brasília, DF,
Brasil. Doutor em Direito. E-mail:
fabianohpeixoto@gmail.com

Resumo: O artigo tem como objetivo relatar e analisar referências para pesquisas e aplicações envolvendo Inteligência Artificial (IA) e Direito, especialmente no que diz respeito a alguns riscos de ampliação de desigualdades quando se trata de temas de imigração e refugiados. Com a utilização de um referencial canadense, pela maior transparência e com relatório analítico, busca-se alcançar indutivamente reflexões sobre o problema de recepção migratória como elemento de aprofundamento de desigualdades globais sob o impacto de sistemas de inteligência artificial.

Palavras-chave: Desigualdades Globais. Inteligência Artificial. Imigração.

Abstract: This paper aims to report and analyze references for research and applications involving Artificial Intelligence (AI) and Law (AIL), especially with regard to the risks of widening global inequalities when it comes to issues of immigration and refugees. Using a Canadian framework, for greater transparency and with an analytical report, we seek to inductively reflect on the problem of migratory reception as an element of deepening global inequalities under the impact of artificial intelligence systems.

Keywords: Global Inequities. Artificial Intelligence. Immigration.



DOI: <http://dx.doi.org/10.20912/rdc.v15i37.222>

Recebido em: 02.07.2020

Aceito em: 23.07.2020

Introdução

Com uma suspeita de que o sistema de IA para concessão de vistos no Reino Unido seria tendencioso, como afirmado por grupos de direitos humanos, Henry McDonald¹, do *The Guardian*, apresentou uma ideia do problema migratório associado ao uso de inteligência artificial (IA). Segundo McDonald, ativistas na proteção de direitos humanos (*Foxglove* e o *Joint Council for the Welfare of Immigrants – JCWI*) questionam o uso do algoritmo pelo Ministério do Interior do Reino Unido que faz uma filtragem dos pedidos de vistos. O risco está no uso do algoritmo criando classes de candidatos, acelerando o processo para grupos em função da cor (pessoas brancas seriam encaminhadas para um modo mais veloz). Relata a reportagem uma contradição, feita por um porta-voz do Ministério do Interior afirmando que o processo é guiado pela eficiência e que a ferramenta de IA não é usada para decisão, mas apenas para indicar se haverá mais ou menos controle e que isso está em conformidade com a legislação sobre igualdade vigente. Note-se desde já, na introdução do tema no presente texto, que pelo comentário do porta-voz, ao minimizar a atuação da IA circunscrevendo-a à indicação de menor ou maior controle, permite a observação de que isso não deixa de ser, em si, já uma decisão, não? A discussão parece ser uma questão clara de *machine bias*²³, que tanto coloca em risco benefícios da IA e alimenta os mitos associados à inteligência não humana. A classificação nas cores verde, amarelo e vermelho, com níveis proporcionais de controle, que acaba oferecendo agilidade no embarque para pessoas brancas é uma típica situação de *bias*. Por mais que haja a possibilidade de revisão humana, a inclusão em um sistema de controle vermelho aumenta a perspectiva menos favorável do solicitante, conforme observa a reportagem. O JCWI também argumenta na reportagem que a versão algorítmica é uma modernização de um sistema que dizia respeito aos solicitantes ciganos, que teriam recebido tratamento discriminatório. O próprio Reino Unido, à época (Câmara dos Lordes – 2005), já havia reconhecido que o estereótipo de cigano, como menos propenso a ser turista genuíno, era ilegal. A reportagem mostrou um problema, mas seria um problema da IA ou um desvio na conduta humana captado (involuntariamente ou propositalmente) pelo sistema de reprodução da atividade cognitiva humana que é, aliás, uma definição possível para a própria IA?

1 MCDONALD, Henry. AI system for granting UK visas is biased, rights groups claim. Immigrant rights campaigners bring legal challenge to Home Office on algorithm that streams visa applicants. *The Guardian*, 29 de outubro de 2019. Disponível em <https://www.theguardian.com/uk-news/2019/oct/29/ai-system-for-granting-uk-visas-is-biased-rights-groups-claim>. Último acesso em 14/03/2020.

2 ELIAS, Paulo Sá. Algoritmos, Inteligência Artificial e o Direito. *Conjur*, novembro, 2017. Disponível em: <https://www.conjur.com.br/dl/algoritmos-inteligencia-artificial.pdf>. Último acesso em 02/3/2020.

3 POLONSKI, Vyacheslav. **Mitigating algorithmic bias in predictive justice: 4 design principles for AI fairness**: Algorithms are being used to convict criminals and decide jail time. We need to make sure they are fair. 2018. Disponível em <https://towardsdatascience.com/mitigating-algorithmic-bias-in-predictive-justice-ux-design-principles-for-ai-fairness-machine-learning-d2227ce28099>. Último acesso em 02/03/2020.

Em uma situação completamente diversa, em fevereiro de 2020, Nita Bhalla⁴ apresentou um artigo apontando que em Uganda, médicos estão aplicando IA para diminuir a mortalidade materna após o nascimento dos seus bebês. O sistema, relativamente simples, acompanha, monitora e transmite os dados das pacientes, pela internet (solução já incorporada no dispositivo), tais como frequência respiratória, pulsação, níveis de saturação de oxigênio e pressão arterial a um *dataset*, em que os algoritmos detectam casos de risco e alertam aos médicos. Em um ambiente de recursos limitados, em que milhares de mulheres morrem de causas evitáveis relacionadas à gravidez e parto e que as instalações médicas são deficitárias isso pode ser extremamente favorável à proteção da vida. Celeridade burocrática, redução de intervenção humana em atividades repetitivas, demoradas e sujeitas a erros não intencionais humanos, precisão, eficiência podem estar associadas à concretização de direitos, dos mais nobres, como a proteção à saúde, integridade, liberdade, mobilidade, dignidade e a própria proteção da vida. Olly Buston, CEO da *Future Advocacy* mencionou⁵ que há um futuro alternativo para a IA propor soluções para problemas de crescimento econômico, prosperidade compartilhada e cumprimento dos direitos humanos e que serão escolhas de empresas, governos e pessoas que determinarão o caminho que a humanidade seguirá.

Esse texto tem justamente como objetivo refletir sobre as necessárias referências para a pesquisa e, mais especificamente, o uso de aplicações de inteligência artificial, com efeitos jurídicos, no que diz respeito a questões migratórias e sistemas de refugiados. No tema, a preocupação central será em alguma medida pensar sobre os riscos de se ampliar desigualdades globais quando se trata dos temas de imigração e sistemas para refugiados e a partir deles, como oportunidade, estabelecer referenciais para sistemas de IA mais transparentes, equitativos e justos. A abordagem que será feita parte da premissa da alta complexidade que envolve a incidência de automação tecnológica, especialmente inteligência artificial (IA), em temas de uma ciência social aplicada. A combinação do raciocínio jurídico e de demandas éticas no direito, combinadas com o raciocínio característicos das áreas de engenharia de software e ciência da computação torna ainda mais delicado o problema de *bias*, e o risco de reforço de infraestrutura para a afirmação do direito como o instrumento de sustentação de desigualdades e não de seu combate. As decisões, com base e classificações, triagens e padrões suportados por sistemas de IA podem, sem o devido resguardo metodológico, atendimento de princípios éticos robustos e mecanismos de segurança e *accountability* sustentar um modelo (com força normativa) de aprofundamento de desigualdades e supressão de direitos. Em contrapartida, pela

4 BHALLA, Nita. **Ugandan medics deploy AI to stop women dying after childbirth**. Thomson Reuters Foundation. News. Disponível em <https://news.trust.org/item/20200131160316-lp5sv/>. Último acesso em 13/3/2020.

5 In: Business&Human Rights Resource Centre. **Technology and Human Rights: Artificial Intelligence**. Disponível em <https://www.business-humanrights.org/en/technology-and-human-rights-artificial-intelligence>. Último acesso em 13/3/2020.

inserção na chamada “Era da IA”, os sistemas têm o potencial de acelerar mecanismos de proteção e mobilidade, muitas vezes vitais e decisivos para proteção e concretização de direitos humanos na temática.

1 Um cenário no controle migratório e de refugiados no Canadá

1.1 A aplicação de IA no sistema migratório canadense

O *The Citizen Lab* em conjunto com o *International Human Rights Program* da Faculdade de Direito da Universidade de Toronto publicaram uma pesquisa de Petra Molnar e Lex Gill⁶ chamada *Bots at the gate: a Human Rights analysis of automated decision-making in Canada's immigration and refugee system*. Neste extenso trabalho foi apresentada uma visão sobre o a utilização de IA no sistema migratório canadense e as questões de análise preditiva, taxonomia, impactos nos direitos humanos, incluindo direitos de igualdade, liberdade, associação, religião, expressão, movimento e mobilidade, privacidade, segurança e à um processo administrativo devido e justo. É, portanto, uma abordagem ampla de um problema intensificado pela aplicação de avanços tecnológicos nos movimentos migratórios espontâneos e forçados. É também um exemplo importante pelo fato do Canadá ser um dos países com um maior nível de regulação do uso e do desenvolvimento de IA e um dos mais destacados em termos de documentos⁷ com preocupações éticas no desenvolvimento e uso de sistemas inteligentes.

A grande preocupação que motivou o trabalho foi, portanto, com os impactos concretos que a tomada de decisão automatizada pode causar nos sistemas de gestão da imigração e de refugiados por um aumento de discricionariedade em um ambiente formado por muitas comunidades vulneráveis e com poucos recursos para qualquer tipo de questionamento ou enfrentamento. Foi identificado um fator crítico com a proteção

6 MOLNAR, Petra; GILL, Lex. **Bots at the gate: a Human Rights analysis of automated decision-making in Canada's immigration and refugee system**. International Human Rights Program (Faculty of Law, University of Toronto) and the Citizen Lab (Munk School of Global Affairs and Public Policy, University of Toronto). Universidade de Toronto. Toronto. Canadá .2018. Disponível em <https://citizenlab.ca/2018/09/bots-at-the-gate-human-rights-analysis-automated-decision-making-in-canadas-immigration-refugee-system/>. Último acesso em 14/3/2020.

7 Treasury Board Directive on the Use of Machine Learning for Decision-Making. Disponível em <https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592>. Último acesso em 14 de janeiro de 2020. CANADÁ, *Institute for Competitiveness & Prosperity*. *From Prediction to Reality. Ontario's AI opportunity. Working Paper 32, June. 2018*. Disponível em https://www.competeprosper.ca/uploads/2018_From_prediction_to_reality_Ontarios_AI_opportunity.pdf. Último acesso em 18/02/2020. *Montreal Declaration for a responsible development of artificial intelligence. 2018*. Disponível em https://docs.wixstatic.com/ugd/ebc3a3_bfd718945e0945718910cef164f97427.pdf. Último acesso em 18/02/2020. CANADÁ, *Artificial intelligence and society. Summit of the G7 science academies. 2019*. Disponível em <https://rsc-src.ca/sites/default/files/Artificial%20intelligence%20and%20society%20G7%202019.pdf>. Último acesso em 18/02/2020. CANADÁ, *Institute for Competitiveness & Prosperity*. *From Prediction to Reality. Ontario's AI opportunity. Working Paper 32, June. 2018*. Disponível em https://www.competeprosper.ca/uploads/2018_From_prediction_to_reality_Ontarios_AI_opportunity.pdf. Último acesso em 18/02/2020.

menos robusta de direitos humanos e as dificuldades de defesa desses direitos. O relatório não se posicionou de forma repulsiva à própria tecnologia, mas ao uso irresponsável dela, podendo exacerbar disparidades. Identificou-se que o governo canadense, desde 2014 ampliou o uso de tecnologia e está desenvolvendo um sistema de análise preditiva para automatizar atividades até então conduzidas por funcionários de imigração. O relatório também afirma que o governo canadense, de forma muito discreta, está amadurecendo a possível utilização de solução de IA para tomada de decisão e para avaliação de aceitação ou não de imigração, inclusive com dispositivo de classificação de risco para “pré-remoções” e utilização de *datasets* com alimentação sobre pessoas vulneráveis, vítimas de violência ou guerra ou que recorreram a sistemas de proteção humanitária, ou, ainda, sistemas compassivos.

A primeira questão que surge no relatório⁸ é que a natureza sutil e complexa de reivindicações de refugiados e imigrantes pode ser perdida, com a consequente violação de direitos humanos protegidos internacionalmente, concretizados sob a forma de preconceitos, discriminações, violações de privacidade, entre outras. O tema analítico proposto no trabalho tem fundamento jurídico, já que as várias facetas do uso de sistemas de decisão automatizados podem atingir direitos humanos, incluindo direitos à igualdade e à não discriminação; liberdade de movimento, expressão, religião e associação; privacidade, vida, liberdade e segurança das pessoas, como dito, e também despertam questões de direito constitucional e administrativo; acesso à justiça, responsabilidade pública e privada, capacidade de gestão pública e governamental, bem como outros impactos globais.

A pesquisa⁹ identifica premissas para tornar o problema da discriminação mais marcante ao estabelecer como lente o momento de maior segurança e controle de fronteiras e medidas de triagem sobre migração global em uma época onde há um considerável aumento global da criminalização da migração e aumento da xenofobia. Afirma que o Canadá tem obrigações domésticas e internacionais claras com o respeito e proteção dos direitos humanos e que cabe aos políticos, funcionários públicos, tecnólogos e engenheiros assim como advogados, sociedade civil e universidade adotar uma ampla visão crítica dos impactos reais do uso de tecnologia sobre a vida humana.

Algumas situações concretas de aplicação de IA relacionadas à *bias*, foram trazidas, como modelos de previsão de riscos de reincidência; automatização de identificação de fraudes no sistema de proteção ao desemprego e predição de pontos críticos de crimes futuros para demonstrar o risco de se substituir tribunais administrativos, oficiais de imigração, agentes de fronteira, analistas legais por sistemas de IA¹⁰. A preocupação

8 MOLNAR, Petra. 2018.

9 MOLNAR, Petra. 2018.

10 Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, “Machine Bias,” ProPublica (23 Ma y 2016) <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> Último acesso em

externada no relatoria não é sobre a distinção entre assistência ou tomada de decisão em si, mas os efeitos da inserção de IA no controle de imigração e refugiados canadense como um todo. E, segundo o relatório, a preocupação é maior ainda porque comunidades vulneráveis e com poucos recursos como os não cidadãos geralmente têm acesso a proteções menos robustas.

O relatório também reconhece que em resposta a questões de migração, até com a melhor das intenções, formuladores de políticas, pela ânsia de usar novas tecnologias como uma solução mais rápida para questões complexas e muitas vezes intratáveis. Sugere a pesquisa crítica, empírica e orientada à proteção absoluta de direitos como um importante contrapesos. Assim, a pesquisa se posiciona claramente sobre o problema: a questão não é como usar a nova tecnologia para consolidar problemas antigos, mas para entender melhor como pode-se usar essa oportunidade para imaginar e projetar sistemas mais transparentes, equitativos e justos.

1.1.1 Atividades feitas pela IA

Desde 2014 o IRCC (Imigração, Refugiados e Cidadania do Canadá) orienta o desenvolvimento de uma análise preditiva para automatizar decisões na avaliação de imigrantes, identificando o mérito de um pedido, bem como, evidenciando possíveis sinais de alerta. O alcance de sistemas de decisão automatizado pode se dar pela execução de tarefas como aponta o relatório¹¹:

Tabela 1. Atividades de IA no sistema de migração e refugiados canadense¹²:

Atividades de IA no sistema de migração e refugiados canadense	
1	Classificar casos, aplicativos ou indivíduos para triagem (por exemplo, em termos de risco, prioridade ou complexidade).
2	Gerar pontuações, avaliações de probabilidade e outros indicadores para serem considerados como fatores de apoio ao raciocínio de um tomador de decisão humano.
3	Identificar ou “sinalizar” certos casos para revisão ou investigação humana.
4	Fornecer recomendações gerais sobre se um aplicativo deve ser aprovado.
5	Prestar a decisão administrativa completa.

15/3/2020 ; Robert N. Charette, “Michigan’s MiDAS Unemployment System: Algorithm Alchemy Created Lead, Not Gold,” IEEE Spectrum (24 January 2018) <https://spectrum.ieee.org/riskfactor/computing/software/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold>; Último acesso em 15/3/2020; See generally Andrew Guthrie Ferguson, *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement* (New York: NYU Press, 2017).

11 MOLNAR, Petra. 2018, p. 3-14.

12 MOLNAR, Petra. 2018, p. 3.

Segundo o relatório, nessas atividades, na medida que se tornem normalizadas e integradas ao sistema canadense de imigração e refugiados – com estão em até certo ponto sendo utilizadas e ampliadas, é indispensável o fundamento em princípios de uso transparente, responsável, justo e, especialmente, o respeito aos direitos humanos. Há uma necessidade urgente, para o relatório, do engajamento acadêmico e da sociedade civil na questão do uso adequado da IA. A escala dos impactos também é apontada no relatório, trazendo números de mais de 50.000 solicitações no ano de 2017. Há um universo de milhares de pessoas afetadas (segurança física, direitos humanos, meios de subsistência, etc.) em seus movimentos migratórios em função de estudos, trabalhos e novas oportunidades de vida. Para refugiados, são ainda mais graves as consequências de uma demora excessiva, custos de procedimentos extras, detenção, separação de familiares ou rejeição de uma proteção, de maneira errônea em função de algum *bias* (raça, religião, nacionalidade, participação de um grupo social específicos ou opinião política, ou histórico de proteção humanitária ou atividade compassiva). Tudo isso pode desencadear em ameaças a situações de tortura, tratamento ou punição cruel e desumana ou risco à vida ocasionado pela origem do solicitante de refúgio. Portanto, as salvaguardas das normas de proteção migratória e de refugiados devem abranger também a aplicação de sistemas de IA.

Tabela 2. Números migratórios¹³:

Atividades de IA no sistema de migração e refugiados canadense	
2017	50.000 solicitações de refugiados.
2018	340.000 novos residentes permanentes (projeção).
2020	340.000 novos residentes permanentes (projeção).
2016	266.000 estudantes com permissão de estudo internacional
2016	10.000 licenças para residentes temporários, 78.500 autorizações de trabalho temporário, 1,3 milhão de autorizações para visitantes e 2,5 milhões de autorizações eletrônicas de viagem (ETA)

De uma das referências do relatório é possível acessar um documento com descrições específicas de atividades de uma chamada para solução de IA (B8607-180311/A), publicada para fechamento em junho de 2018, com o seguinte *background*: tendo como *datasets* os registros do RFI (pedido de informações); do IRCC (Imigração, Refugiados e Cidadania do Canadá); do ESDC (Emprego e Desenvolvimento Social do Canadá) e do DOJ (Departamento de Justiça), desenvolver uma benefício de IA, utilizando *machine learning*, *predictive analytics* e *support decisions*, para modelos mais eficientes com considerações legais, estratégicas para litígios e suporte ao governo canadense.

13 MOLNAR, Petra. 2018, p. 4

Basicamente a solução deve atender às preocupações canadenses sobre pesquisa jurídica (assessoria jurídica) nas decisões, como avaliação de riscos legais e previsão de resultados em litígios futuros.

Essa identificação de objetivos é sintomática, pois indica uma tendência de orientação de ação estatal pela previsão de litígio e antecipação de riscos, não necessariamente pela interpretação de direitos. Trata-se de orientação típica da iniciativa privada na perspectiva custo/lucro.

Outra constatação desta chamada é a busca por uma solução de *machine learning* que possa ser expansível para usuários tomadores de decisão do *front-end*. Além disso, a ideia apresentada busca o compartilhamento de informações e parâmetros entre o IRCC, ESDC e DOJ. O contexto de compartilhamento com *datasets* de imigração e refugiados com o aparato de segurança nacional do Canadá também demonstrou, segundo o relatório, inconsistências. Avaliações de segurança, associadas a atividades criminais e sob investigação pelo CSIS (Serviço de Inteligência de Segurança do Canadá) retroalimentam o sistema de admissão. O exemplo marcante citado foi o ocorrido na RCMP (*Royal Canadian Mounted Police*), que enfrentou severas críticas por incluir no seu guia de entrevista na fronteira, perfis religiosos e étnicos de migrantes. A RCMP coletou em um questionário, aplicado a solicitantes de asilo, com perguntas com estereótipo islamofóbico, solicitando respostas sobre crenças políticas e religiosa, percepção do indivíduo sobre mulheres que não usam *hijab*, opinião sobre o Estado Islâmico e Talibã, bem como a quantidade de vezes que o solicitante orava por dia. Não havia no questionário nenhuma pergunta sobre outras vertentes religiosas ou sobre outros grupos terroristas. Essas respostas foram inseridas em um banco de dados e compartilhado como outros serviços governamentais.

Portanto, este documento já em 2018 indicava uma tendência, de ampliação de *datasets* com o compartilhamento de informações e a possível composição deles na formação da arquitetura desviada de *machine learning*, com dados sobre fraude e atribuição de pesos para formular uma recomendação se o candidato deve ser aceito ou recusado. Sob o ponto de vista da existência de dados sobre populações vulneráveis (que são associadas a custos estatais nos seus processos de migração, refúgio e fruição de direitos básicos), isso pode gerar barreiras fortes no sentido de respeito aos direitos humanos, afetando severamente a preocupação básica do relatório de usar a nova tecnologia como oportunidade para imaginar e projetar sistemas mais transparentes, equitativos e justos.

Uma fundamental observação para o tema foi apresentada no relatório ao mencionar o funcionamento dos sistemas de decisão autônomos e pode fundamentar o caminho para a confiabilidade dos modelos: eles têm como base um sistema de considerações com referenciais de eficiência correlacionados com um conjunto de dados. Portanto, a estrutura de correlação é formada por um arranjo variável, que pode estar metodologicamente estruturado para perpetuar ou exacerbar padrões de preconceito ou

discriminação, bem como para detectá-los, identificá-los e enfrentá-los. O relatório se apresenta cético com relação a essa possibilidade, pois fundamenta que todas as opções tecnológicas (sobre o que contar, quem conta e por que) têm uma dimensão inerentemente política.

2.1.2 recomendações do relatório para a redução de riscos:

O relatório explica bases conceituais para a IA e, nisso, ajuda a desmistificar a função dos algoritmos. Algoritmos de *machine learning* permitem, como uma receita, o treinamento utilizando um grande *corpus* de dados para classificar e generalizar a partir de exemplos no conjunto de treinamento. Portanto, disso é possível perceber que os desvios não são propriamente dos algoritmos, mas no conjunto de dados selecionados para o treinamento. É perfeitamente possível, sob o ponto de vista metodológico e de standards de desenvolvimento se prever, mapear e selecionar dados que apontem para os principais desvios, preconceitos e situações indesejadas, bem como indicar etapas de controle para anotação e retreinamento quando ocorrer algum desvio indesejado ou identificado. São cuidados e etapas exequíveis que devem estar na arquitetura de uma solução sólida de IA.

O trabalho apresentado indica comportamentos para a minimização dos riscos de aumento de desigualdade na forma de recomendações ao governo canadense:

Tabela 3. Recomendações do Relatório para o governo canadense¹⁴:

	Recomendação
1	Publique um relatório completo e detalhado, a ser mantido continuamente, de todos os sistemas de decisão automatizados atualmente em uso no sistema de imigração e refugiados do Canadá, incluindo informações detalhadas e específicas sobre cada sistema.
2	Congele todos os esforços para adquirir, desenvolver ou adotar qualquer nova tecnologia de sistema de decisão automatizada até que os sistemas existentes obedeçam totalmente a uma norma ou diretiva de todo o governo que rege o uso responsável dessas tecnologias.
3	Adote uma norma ou diretiva vinculativa em todo o governo para o uso de sistemas de decisão automatizados, que deve ser aplicada a todos os novos sistemas de decisão automatizados, bem como aos atualmente em uso pelo governo federal.
4	Estabeleça um órgão independente, com o poder de se envolver em todos os aspectos da supervisão e revisar todo o uso de sistemas de decisão automatizados pelo governo federal.
5	Crie uma metodologia racional, transparente e pública para determinar os tipos de processos e sistemas administrativos que são apropriados para o uso experimental de tecnologias automatizadas de sistemas de decisão e quais não são.
6	Comprometer-se a criar código fonte completo para todos os sistemas de decisão automatizados do governo federal - independentemente de serem desenvolvidos internamente ou pelo setor privado - por padrão público e de código aberto, sujeito apenas a exceções limitadas por razões de privacidade e segurança nacional.

¹⁴ MOLNAR, Petra. 2018, p. 2.

- | | |
|---|--|
| 7 | Lance uma força-tarefa federal que traga as principais partes interessadas do governo, juntamente com o meio acadêmico e a sociedade civil, para entender melhor os impactos atuais e prospectivos das tecnologias de sistemas de decisão automatizados nos direitos humanos e no interesse público de maneira mais ampla. |
|---|--|

O relatório ainda aponta itens que devem ser garantidos aos refugiados:

- 1) Direito de ter uma entrevista pessoal;
- 2) Direito de receber a divulgação completa das informações consideradas no seu caso;
- 3) Direito ao tratamento ético e com justificção para decisão automatizada baseada em critério de correção.

3 Redução dos riscos para sistemas de decisão automatizada:

3.1 Critérios claros

Antes de ingressar propriamente nas possibilidades para redução de riscos no desenvolvimento e uso de IA, retoma-se o relatório de referência, que apresenta um quadro muito bem estruturado com as perguntas para se identificar os impactos e riscos de um sistema autônomo de decisão:

Tabela 4. Perguntas par identificação dos potenciais impactos e riscos dos Automated Decision Systems¹⁵:

	Pergunta
1	Quais os critérios que serão usados para definir e avaliar fraude ou informação falsa/adulterada?
2	Que tipo de dados ou evidências serão coletados e alimentados no sistema automatizado?
3	Quem terá acesso a essas informações e como elas serão compartilhadas com outros departamentos?
4	O que o governo considera margem de erro aceitável para esses sistemas?
5	Quais serão os motivos de apelação ou reparação se um sistema automatizado fizer ou apoiar uma declaração de inadmissibilidade contra você?

A tabela 4 apresenta elevado detalhamento dos possíveis impactos e necessidades para a organização metodológica do desenvolvimento de um sistema de inteligência artificial. to

¹⁵ MOLNAR, Petra. 2018, p.5.

3.2 Organização metodológica do desenvolvimento

Segundo Cid¹⁶ (2020, p. 78) “a coexistência de inteligência natural e artificial está tendo e terá consequências para a sociedade e seus cidadãos em todas as esferas de sua existência” e completa com a afirmação de que “a pesquisa acelerada e incessante e a aplicação da IA destacam as lacunas e deficiências das teorias legais ‘analógicas’ tradicionais para responder às necessidades e demandas do atual contexto tecnológico”.

O desenvolvimento de um sistema de IA em área de afetação do direito impõe um trabalho formado por equipes de tecnologia (por exemplo de ciência da computação, engenharia de software, etc.) e de direito. A realização de um plano de trabalho completo e profundo deve apresentar a definição de objetivos, metodologia e resultados esperados sob a ótica de pesquisa e desenvolvimento. Nos objetivos devem estar presentes os riscos a serem evitados e os mecanismos de proteção e segurança contra *bias*. No projeto também devem estar claras as informações sobre transparência e *accountability*. A metodologia deve considerar que os humanos são muito mais criativos do que os programas de computador que eles escrevem. No entanto, os computadores são mais sistemáticos e menos propensos a erros em tarefas repetitivas que os humanos, ou seja, sob o ponto de vista metodológico deve-se otimizar uma qualidade da IA, observada por Eagle¹⁷ (2004), que é a capacidade de desempenhar um papel útil como busca de diagnósticos e verificações (*checklists*). Para ele a IA pode atuar como *backstop* para o raciocínio humano, prevenindo erros, isto é, aumentando acertos, evitando omissões, potenciais pleitos ou possíveis linhas de argumentação robustas.

3.3 Preparação de datasets

Percebe-se no relatório canadense um recorrente fator crítico para aplicação de *machine learning* ou outras áreas da IA: a necessidade de dados corretos e o risco de dados distorcidos ou “contaminados”. Se o próprio conceito de correção é complexo, pode-se trabalhar com a eliminação de dados identificáveis como incorretos (como o observado na coleta do questionário da *Royal Canadian Mounted Police*, produzindo dados islamofóbicos). Além da confiança nos dados, deve existir uma autenticidade na

16 CID, Isabel-Victoria Lucena. Technological Revolution: an approach to the new Technologies from the perspective of Human Rights. In: Las Fronteras de los Derechos Humanos: problemas, discusión y soluciones. Dykinson eBook. Ramón Luis Soriano Diaz, David Sánchez Rubio, Juan Carlos Suárez Villegas Editores. ISBN 978-84-1424-656-7. Madrid. 2020. p.78. No original: “The coexistence of natural and artificial intelligence is having and will have consequences for society and its citizens in all spheres of their existence. This is an ever-expanding field that drives us to define fundamental questions related with the legislation of this sphere. The accelerated and incessant research and the application of AI highlight the gaps and shortcomings of traditional ‘analogical’ legal theories to respond to the needs and demands of the current technological context.”(tradução livre nossa).

17 EAGLE, Eric Allen. An Introduction to Artificial Intelligence and Legal Reasoning: Using x Talk to Model the Alein Tort Claims Act and Torture Victim Protection Act. Richmond Journal of Law & Technology. Volume XI, Issue 1. 2004. <http://jolt.richmond.edu/jolt-archive/v11i1/article2.pdf>. Último acesso em 05/05/2020.

visão da própria construção, pois pode ser usada para suporte a alguma outra decisão, mesmo que não seja algum tipo de decisão peremptória.

Verificado *a priori*, a formação do *corpus* busca uma proteção contra preconceitos injustos e uma paridade de representação no *dataset*, evitando o risco de sub-representação na fase de treinamento. Essa é uma condição necessária, mas ainda não suficiente para o combate a modelos eventualmente tendenciosos ou preconceituosos.

Para um adequado enfrentamento das tendências e otimização das qualidades benéficas da tecnologia, deve haver transparência e divulgação responsável em torno de sistemas de IA para garantir que as pessoas entendam os resultados baseados em IA e possam questioná-los. Inclui *accountability* e, registro de dados e documentação do design. O modelo de transparência e *accountability* deve, sempre que possível, envolver um processo de certificação de boas práticas a ser ofertado pelos entes envolvidos (academia, indústria e governo). Esse modelo deve estar presente desde o desenvolvimento e envolver as etapas de verificação, validação, controle e segurança do sistema.

A negação ao uso da IA chega a ser ingênua, o pensamento estratégico deve se direcionar para buscar potencializar benefícios potenciais e transparecer os problemas da decisão humana enviesada. No mesmo sentido, a identificação clara de autoridade e competência no âmbito da responsabilidade é um mecanismo eficiente de proteção das qualidades benéficas da velocidade, alcance, transparência lógica e precisão. As pessoas devem exercer níveis adequados de julgamento e permanecerem responsáveis pelo desenvolvimento, implantação, uso e resultados dos sistemas de IA. O modelo de responsabilização deve, sempre que possível, envolver um processo de certificação de boas práticas a ser ofertado pelos entes envolvidos no desenvolvimento e uso de IA (academia, indústria e governo juntos). Se o volume de operações é dado como um fator limitador da responsabilidade, há o contraponto dos necessários mecanismos de segurança e controle e dos espaços de anotações de desvios identificados. A ideia é que o sistema de IA deixe as fragilidades e tendências que são humanas sob luzes.

3.4 parâmetros compatíveis com os direitos humanos

Uma outra potencialidade de sistemas de IA, que não a formulação de sistemas *decision-maker*, está na possibilidade de potencializar sistemas de informação dos mais variados, inclusive os necessários para se formar algo como uma perspectiva quantitativa e qualitativa de imaginário coletivo sobre a necessidade de respeito aos direitos humanos em processos e crises migratórios. Este possível benefício não pode ser desprezado. Ramírez¹⁸ (2020) menciona uma série de fenômenos associados ao incremento digital

18 RAMIREZ, Marta Gil; ROJAS, Ruth Gómez de Travesedo. **Derechos Humanos e Inmigración.** Repercusión del caso 'Aquarius' en Youtube. In: Derechos Humanos desde la Interdisciplinarietà en Ciencias Sociales y Humanidades. Dykinson eBook. David Sánchez Rubio, Sergio Marín-Conejo e Jorge Olvera García Editores. ISBN 978-84-1324-651-2. Madrid. 2020.

de repercussão da temática em redes sociais, substituindo (suplantando) os meios de disseminação de informações tradicionais. Observa que “[...] depois de confirmar uma evidente falta de pesquisa que aborde o papel que as redes sociais desempenham hoje na construção do imaginário coletivo sobre o fenômeno migratório em geral e sobre a figura do migrante em particular [...]”, indica que “[...] parece necessário iniciar uma linha de pesquisa nesse sentido, negligenciada, no momento, pela comunidade científica.¹⁹”

O uso de IA para informação relevante e derrubada de *deepfake news* sobre o tema é estratégico e fundamental para a concretização do respeito à direitos humanos, propagação de informação do uso enviesado e pressão social para adequações.

O respeito aos direitos humanos e o combate a já tão acentuada desigualdade global deve impor aos sistemas de IA uma arquitetura de respeito o estado de direito, valores democráticos e a diversidade, que deve incluir salvaguardas apropriadas (por exemplo, possibilitando a intervenção humana para anotação de erros e novo treinamento pelas diretrizes da lealdade e justiça substancial).

Um caminho muito interessante para se chegar a parâmetros parte da delimitação de um grupo de princípios (justiça substancial, republicanismo, precaução, transparência, etc.). Dos princípios, é possível a definição mais clara de propósitos (governança algorítmica, alfabetização digital, inclusão, diversidade e sustentabilidade. Na reunião e confluência dessas duas categorias: princípios e propósitos, obtém-se parâmetros compatíveis para a inteligência artificial no espectro de respeito aos direitos humanos, ou seja, um sistema de inteligência artificial confiável, robusto e respeitoso²⁰.

Como mencionou Cid (2020, p. 88): “A lei em geral e os direitos humanos em particular têm a capacidade de reinventar e dar forma à inovação, definir as necessidades, benefícios e prioridades que derivam do progresso tecnológico [...]” Desta forma é possível “[...] proteger e garantir os direitos dos cidadãos dos riscos perniciosos que podem derivados do uso e aplicação de novas tecnologias.”²¹ Os sistemas de IA podem e

19 RAMÍREZ, Marta Gil; ROJAS, Ruth Gómez de Travesedo. **Derechos Humanos e Inmigración**. 2020. No original: “tras constatar una evidente carencia de investigaciones que aborden al papel que hoy en día juegan las redes sociales en la construcción del imaginario colectivo sobre el fenómeno migratorio en general y sobre la figura del migrante en particular, y dada la especial relevancia que este tipo de plataformas adquieren en el contexto de la globalización, parece necesario iniciar una línea de investigación en este sentido, desatendida, de momento, por la comunidad científica.” (tradução livre nossa).

20 HARTMANN PEIXOTO, Fabiano. *Inteligência Artificial e Direito: convergência ética e estratégica*. Ed. Alteridade. Curitiba. 2020.p. 163.

21 CID, Isabel-Victoria Lucena. *Technological Revolution: an approach to the new Technologies from the perspective of Human Rights*. In: *Las Fronteras de los Derechos Humanos: problemas, discusión y soluciones*. Dykinson eBook. Ramón Luis Soriano Díaz, David Sánchez Rubio, Juan Carlos Suárez Villegas Editores. ISBN 978-84-1424-656-7. Madrid. 2020 p. 88. No original: “The Law/Technology and Human Rights/New Technologies pairings present an unavoidable relationship in the current context. Law in general and human rights in particular have the capacity to reinvent and give shape to innovation, to define the needs, benefits and priorities that derive from technological progress, and in turn to protect and guarantee the rights of citizens from the pernicious risks that might be derived from the use and application of new technologies.” (tradução livre nossa).

devem ser projetados e operados de modo a serem compatíveis com os ideais de dignidade humana, direitos, liberdades e diversidade e o desenvolvimento e uso da IA devem estar sujeitas ao permanente debate democrático (pela transparência e controle). Da mesma forma que o serviço de controle migratório e atendimento inicial ao refugiado é prestado pelo humano (servidor público ou em desempenho de uma função pública) passível de controle e responsabilização, o sistema artificial (que é uma reprodução de padrões humanos) também pode e deve estar associado a um sistema de responsabilidade pessoal.

Assim, os sistemas de IA devem ser desenvolvidos e aplicados no direito com a finalidade muito clara de beneficiar as pessoas e o nosso ecossistema global, impulsionando o crescimento inclusivo e sustentável e o bem-estar, melhorando as condições gerais de trabalho, saúde e vida das pessoas em geral. A restrição de bem-estar individual, com implica a restrição a direitos migratórios ou refúgio deve estar clara, transparente, justificada e auditável, amparada por critérios de justiça substancial.

Conclusões

Uma constatação para a chamada consistência da IA no direito é a ampliação do que se considera bom desempenho para se aproximar da ideia de desempenho justo (um conceito extremamente complexo), mas que pode nos associar a elementos de boa-fé, lealdade, solidariedade, responsabilidade, ética, concretização de direitos fundamentais, etc.

De uma forma tradicional, com essas tecnologias é possível incrementar a celeridade e acurácia na análise de uma quantidade expressiva de processos e situações, contemplando princípios relevantes de celeridade, eficiência e economia.

Contudo, acurácia e celeridade não podem ser a métricas-fim ou fundamento raiz da utilização da IA, mas deve existir uma associação com a sustentabilidade, a inclusão e proteção da diversidade, a solidariedade e a equidade para a substancial definição de sucesso para a IA.

O presente texto buscou, a partir de um relatório e um referencial canadense, observar os riscos e apontar possíveis caminhos para o fenômeno do desenvolvimento e uso de sistemas de IA no apoio a decisões nos processos de imigração e sistemas de refugiados.

Há um aumento considerável da oferta de soluções da IA e os Estados estão exercendo a opção pelo uso de IA. Nesse cenário, é importante refletir – apontando já o uso indevido, para buscar-se uma oportunidade do uso benéfico da tecnologia, para contribuir com a aceleração e diminuição dos riscos aos direitos humanos de comunidades vulneráveis.

A primeira parte do texto buscou relatar as considerações do relatório produzido pela Universidade de Toronto, apontando diversas deficiências e impactos no uso da IA, que causam ainda mais preocupação quando associadas a grande quantidade de documentos, muitos deles oficiais, do Canadá no sentido de uso da IA para o bem-estar e proteção do ser humano.

O texto também apresenta recomendações para desenvolvimento e uso de IA nesse contexto de inevitabilidade, para se estabelecer sistemas mais claros, transparentes e auditáveis, potencializando a característica da IA de jogar “luzes” às condutas enviesadas humanas que promovem um agravamento de desigualdades globais.

Referências

CANADÁ, Institute for Competitiveness & Prosperity. **From Prediction to Reality.** Ontário's AI opportunity. Working Paper 32, June. 2018. Disponível em https://www.competeprosper.ca/uploads/2018_From_prediction_to_reality_Ontarios_AI_opportunity.pdf. Último acesso em 18/02/2020.

CANADÁ, **Montreal Declaration for a responsible development of artificial intelligence.** 2018. Disponível em https://docs.wixstatic.com/ugd/ebc3a3_bfd718945e0945718910cef164f97427.pdf. Último acesso em 18/02/2020.

CANADÁ. G7 Multistakeholder Conference on AI. Theme 1: **AI for Society: Inclusion in AI Development and Deployment.** Discussion Paper for Breakout Session. Disponível em [https://www.ic.gc.ca/eic/site/133.nsf/vwapj/1_Discussion_Paper_-_AI_for_Society_EN.pdf/\\$FILE/1_Discussion_Paper_-_AI_for_Society_EN.pdf](https://www.ic.gc.ca/eic/site/133.nsf/vwapj/1_Discussion_Paper_-_AI_for_Society_EN.pdf/$FILE/1_Discussion_Paper_-_AI_for_Society_EN.pdf). Último acesso em 22 de fevereiro de 2020.

CANADÁ, Summit of the G7 science academies. **Artificial intelligence and society.** 2019. Disponível em <https://rsc-src.ca/sites/default/files/Artificial%20intelligence%20and%20society%20G7%202019.pdf>. Último acesso em 18/02/2020.

CID, Isabel-Victoria Lucena. Technological Revolution: an approach to the new Technologies from the perspective of Human Rights. In: **Las Fronteras de los Derechos Humanos: problemas, discusión y soluciones.** Dykinson eBook. Ramón Luis Soriano Diaz, David Sánchez Rubio, Juan Carlos Suárez Villegas Editores. ISBN 978-84-1424-656-7. Madrid. 2020.

CHEN, Stephen. **Artificial intelligence, immune to fear or favour, is helping to make China's foreign policy.** South China Morning. Post (30 July 2018) Disponível e <https://www.scmp.com/news/china/society/article/2157223/artificial-intelligence-immune-fear-or-favour-helping-make-chinas>. Último acesso em 18/02/2020.

ELIAS, Paulo Sá. **Algoritmos, Inteligência Artificial e o Direito**. Conjur, novembro, 2017. Disponível em: <https://www.conjur.com.br/dl/algoritmos-inteligencia-artificial.pdf>. Último acesso em 02/3/2020.

GAON, Aviv; STEDMAN, Ian. **A Call to Action: moving forward with the governance of artificial intelligence in Canadá**. Disponível em https://www.researchgate.net/publication/334538357_A_Call_to_Action_Moving_Forward_with_the_Governance_of_Artificial_Intelligence_in_Canada. Último acesso em 18/02/2020.

HARTMANN PEIXOTO, Fabiano. **Inteligência Artificial e Direito: convergência ética e estratégica**. Ed. Alteridade. Curitiba. 2020.p. 163.

HUBBARD, F. **Sophisticated Robots: Balancing Liability, Regulation, and Innovation**. Disponível em https://scholarcommons.sc.edu/law_facpub. Último acesso em 02/3/2020.

MOLNAR, Petra; GILL, Lex. **Bots at the gate: a Human Rights analysis of automated decision-making in Canada's immigration and refugee system**. International Human Rights Program (Faculty of Law, University of Toronto) and the Citizen Lab (Munk School of Global Affairs and Public Policy, University of Toronto). Universidade de Toronto. Toronto. Canadá .2018. Disponível em <https://citizenlab.ca/2018/09/bots-at-the-gate-human-rights-analysis-automated-decision-making-in-canadas-immigration-refugee-system/>. Último acesso em 14/3/2020.

POLONSKI, Vyacheslav. **Mitigating algorithmic bias in predictive justice: 4 design principles for AI fairness**: Algorithms are being used to convict criminals and decide jail time. We need to make sure they are fair. 2018. Disponível em <https://towardsdatascience.com/mitigating-algorithmic-bias-in-predictive-justice-ux-design-principles-for-ai-fairness-machine-learning-d2227ce28099>. Último acesso em 02/03/2020.

RAMIREZ, Marta Gil; ROJAS, Ruth Gómez de Travesedo. **Derechos Humanos e Inmigración**. Repercusión del caso 'Aquarius' en Youtube. In: Derechos Humanos desde la Interdisciplinariedad en Ciencias Sociales y Humanidades. Dykinson eBook. David Sánchez Rubio, Sergio Marín-Conejo e Jorge Olvera García Editores. ISBN 978-84-1324-651-2. Madrid. 2020.

SURDEN, Harry, **The Ethics of Artificial Intelligence in Law: Basic Questions** (August 22, 2019). Forthcoming chapter in Oxford Handbook of Ethics of AI, 2020; U of Colorado Law Legal Studies Research Paper No. 19-29. Available at SSRN: <https://ssrn.com/abstract=3441303>. Último acesso em 02/03/2020.