

GOVERNANÇA DE SISTEMAS DE IA: **UM OLHAR PARA A POPULAÇÃO LGBTQIAPN+**



SUMÁRIO

APRESENTAÇÃO	03
IA E POPULAÇÃO LGBTQIAPN+: QUAL É A DIFERENÇA QUANDO FALAMOS EM RISCOS?	04
COMO ORGANIZAÇÕES PODEM LIDAR COM RISCOS GERADOS POR SISTEMAS DE IA?	05
COMO A GESTÃO DE RISCOS PODE SER CONTEXTUALIZADA À REALIDADE DA POPULAÇÃO LGBTQIAPN+?	06
GOVERNANÇA INCLUSIVA: MEDIDAS PRÁTICAS PARA MITIGAÇÃO DE RISCOS À POPULAÇÃO LGBTQIAPN+	07



APRESENTAÇÃO

O mês de junho é globalmente reconhecido como Mês do Orgulho LGBTQIAPN+, sendo dedicado à visibilidade, à celebração da diversidade e à luta pelos direitos dessa população. Com o objetivo de contribuir às ações que marcam o Mês do Orgulho, elaboramos a presente cartilha, que apresenta dicas práticas para promoção de diversidade e inclusão durante o desenvolvimento e uso de sistemas de inteligência artificial (IA).

Em Prado Vidigal, acreditamos que a IA tem o potencial de gerar benefícios para a sociedade em todos os seus segmentos. Para garantir isso, é necessário partir do reconhecimento e enfrentamento de riscos, criando rotinas de governanças eficazes que promovam a proteção de direitos de grupos vulneráveis, como a população LGBTQIAPN+.

IA E POPULAÇÃO LGBTQIAPN+: QUAL É A DIFERENÇA QUANDO FALAMOS EM RISCOS?

A marginalização, o estigma e a discriminação não se manifestam de forma homogênea. Pessoas podem vivenciar essas realidades de maneiras distintas, influenciadas por fatores como raça, gênero, orientação sexual, identidade regional e outros marcadores sociais de diversidade. Nesse contexto, determinados grupos que historicamente enfrentaram opressões podem estar mais propensos a riscos decorrentes de novas tecnologias, relacionados à perpetuação e amplificação de desigualdades e violências já existentes.

Reconhecer que os riscos decorrentes do desenvolvimento e da implantação de sistemas de IA podem não ser os mesmos para todas as pessoas afetadas é importante para que organizações possam desenvolver ações específicas voltadas a grupos vulneráveis, como é o caso da população LGBTQIAPN+, permitindo a criação de estratégias de mitigação mais eficazes e assertivas.



PARA REFLETIR:

Além de uma exposição intensificada, a população LGBTQIAPN+ pode estar sujeita a diferentes formas de risco. Por exemplo, pesquisadores já demonstraram preocupações com sistemas de IA que buscam “identificar” a orientação sexual de pessoas, especialmente em países onde a homossexualidade e a não-conformidade de gênero são considerados delitos puníveis. Nesses casos, tais sistemas poderiam ser utilizados para perseguição e vigilância, além de perpetuar e reforçar estigmas e falsas concepções que correlacionam a aparência física à orientação sexual e identidade de gênero.¹

¹ TOMASEV, Nenad; MCKEE, Kevin R.; KAY, Jackie; MOHAMED, Shakir. Fairness for Unobserved Characteristics: insights from technological impacts on queer communities. Proceedings Of The 2021 Aai/Acm Conference On Ai, Ethics, And Society, [S.L.], p. 255, 21 jul. 2021. ACM. Disponível clicando aqui.

COMO ORGANIZAÇÕES PODEM LIDAR COM RISCOS GERADOS POR SISTEMAS DE IA?

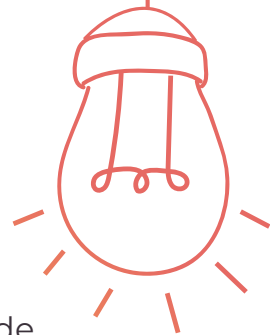
O primeiro passo para as organizações lidarem com riscos gerados por sistemas de IA é identificar e compreender esses riscos. Para isso, as organizações podem recorrer a ferramentas como a taxonomia e a matriz de riscos.

A taxonomia de riscos é uma ferramenta que permite a classificação e categorização dos diferentes riscos que podem ser gerados por sistemas de IA. Uma taxonomia bem estruturada é fundamental para identificar os riscos antecipadamente, o que facilita a implementação de estratégias de mitigação e a adoção de medidas preventivas. Já a **matriz de risco** possibilita a avaliação da probabilidade de ocorrência e do impacto de tais riscos, o que auxilia na priorização de esforços.

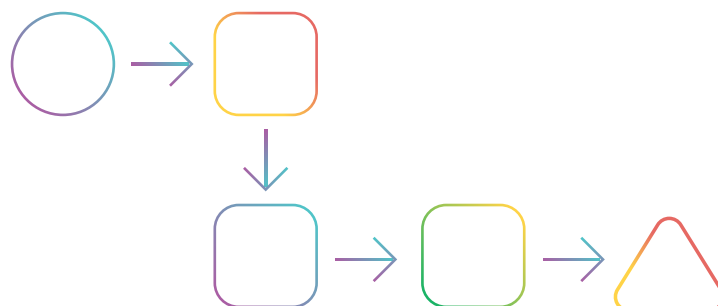
DICA PRÁTICA:

A Taxonomia de Riscos de IA desenvolvida pelo NIST (National Institute of Standards and Technology), que categoriza riscos a partir de atributos técnicos de design, atributos sociotécnicos e valores sociais, pode ser um bom material de referência.

[LEIA AQUI](#)



Juntas, essas ferramentas podem ajudar a identificar, avaliar e priorizar os riscos decorrentes do desenvolvimento ou implementação de sistemas de IA, permitindo que as organizações tomem decisões informadas e minimizem possíveis impactos negativos. Além disso, as taxonomias e matrizes de risco fornecem uma linguagem de governança comum, o que facilita a comunicação entre diferentes áreas e otimiza a compreensão dos riscos e das estratégias para mitigá-los.



COMO A GESTÃO DE RISCOS PODE SER CONTEXTUALIZADA À REALIDADE DA POPULAÇÃO LGBTQIAPN+?

O gerenciamento de riscos precisa ser sensível às particularidades e desafios que a população LGBTQIAPN+ pode enfrentar ao interagir com sistemas de IA. Para ilustrar esse cenário, selecionamos alguns tipos de riscos, a partir da **Taxonomia desenvolvida pelo NIST** e direcionamos um olhar específico para a população LGBTQIAPN+:

CATEGORIAS DE RISCOS	TIPO DE RISCO	CONTEXTUALIZAÇÃO POPULAÇÃO LGBTQIAPN+
Atributos técnicos de design	Acurácia	<ul style="list-style-type: none"> • Sistema apresenta falta de reconhecimento ou representação de identidades dentro do espectro LGBTQIAPN+; • Sistema apresenta conjuntos de dados que excluem ou distorcem a experiência da população LGBTQIAPN+; • Sistema apresenta métodos de coleta de dados que falham em capturar a diversidade das LGBTQIAPN+.
Atributos sociotécnicos	Gestão de vieses	<ul style="list-style-type: none"> • Sistema apresenta resultados discriminatórios ilícitos ou abusivos com base na orientação sexual ou identidade de gênero devido a vieses nos dados de treinamento ou nos próprios modelos; • Sistema perpetua estereótipos negativos sobre pessoas LGBTQIAPN+; • Sistema reproduz microagressões à população LGBTQIAPN+.
Atributos sociotécnicos	Privacidade	<ul style="list-style-type: none"> • Sistema não garante a proteção adequada de dados sensíveis sobre a orientação sexual de pessoas LGBTQIAPN+.

GOVERNANÇA INCLUSIVA: MEDIDAS PRÁTICAS PARA MITIGAÇÃO DE RISCOS À POPULAÇÃO LGBTQIAPN+

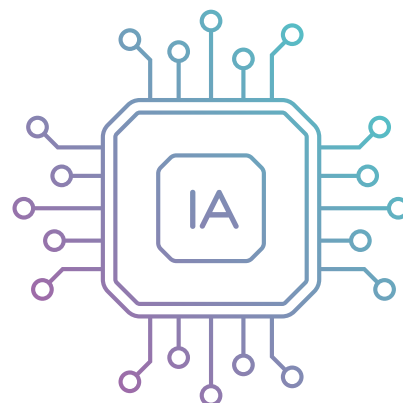
Sabemos que não existe uma solução única e que as rotinas de gestão de riscos devem ser adaptadas à realidade de cada organização, mas é possível citar algumas medidas práticas que organizações podem adotar para mitigação de riscos à população LGBTQIAPN+:

AVALIAÇÃO DE IMPACTO

A Avaliação de Impacto Algorítmico (ou “AIA”) pode ser compreendida como uma ferramenta de accountability utilizada na identificação e mitigação de riscos de um sistema de IA sobre interesses juridicamente e socialmente relevantes.

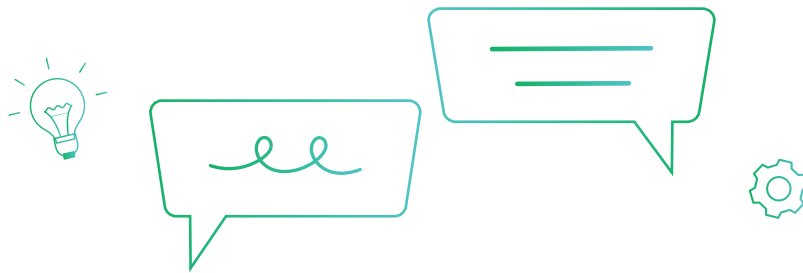
A elaboração deste instrumento tem por objetivo a compreensão de características, limites e potenciais impactos do sistema, buscando, ainda, a mitigação de riscos e a construção de confiança entre as partes interessadas. Conduzir uma avaliação de impacto pode ser uma boa oportunidade para mapear potenciais consequências negativas que sistemas de IA podem ter sobre os direitos da população LGBTQIAPN+.

Essa avaliação pode ser elaborada tanto pela organização responsável pelo desenvolvimento de um sistema de IA, quanto pela organização que implementará, na prática, este sistema. No primeiro caso, a avaliação tem como foco a antecipação de potenciais impactos decorrentes da forma como o sistema foi concebido e programado. Já no segundo caso, a ideia é que a avaliação leve em consideração um contexto real, incluindo pessoas afetadas, ambiente de aplicação, desigualdades e disparidades eventualmente já existentes etc.



COMITÊS DE GOVERNANÇA

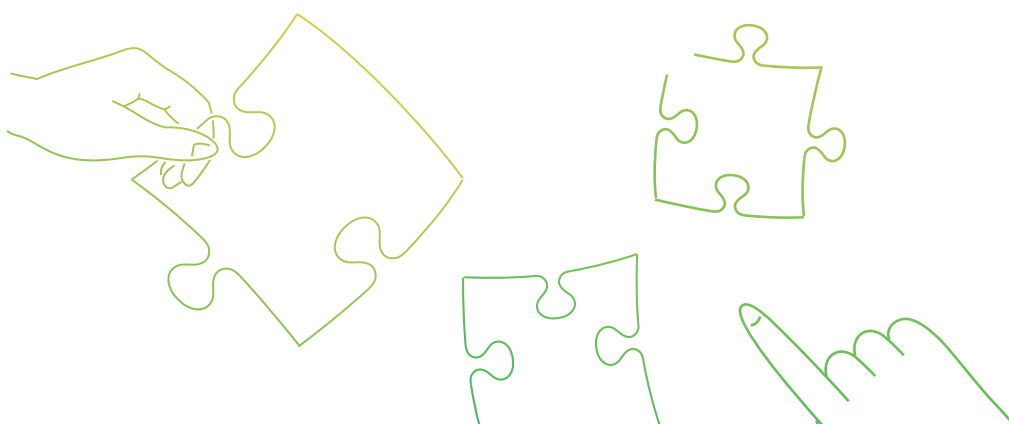
Comitês de governança, como Comitês de Ética e Comitês Multidisciplinares de Governança em IA, garantem que o desenvolvimento e o uso de sistemas de IA estejam alinhados com princípios e diretrizes éticas da organização. Estes comitês funcionam como mecanismos de supervisão, assegurando que as práticas de IA respeitem e promovam valores, como a diversidade e a inclusão.



Para tanto, as organizações podem trabalhar para que a composição dos comitês conte com ampla gama de visões e perspectivas, por exemplo, com representação da população LGBTQIAPN+ por meio de membros fixos ou rotativos, internos ou externos, quando aplicável (por exemplo, especialistas, consultores e membros de grupos afetados). Dessa forma, é possível assegurar que os interesses e necessidades da população LGBTQIAPN+ sejam representados e considerados nas decisões.

INCENTIVO À DIVERSIDADE EM EQUIPES DE DESENVOLVIMENTO

Para além da diversidade de perspectivas em Comitês de Governança, organizações podem promover incentivos à formação de equipes técnicas e de desenvolvimento de sistemas de IA diversas. Essa é uma prática relevante para possibilitar que, desde o momento de concepção dos sistemas, uma ampla gama de perspectivas e experiências sejam consideradas e representadas, gerando soluções mais inovadoras, éticas e robustas.



CONSCIENTIZAÇÃO

Sensibilizar a equipe interna sobre os desafios que novas tecnologias podem trazer para a população LGBTQIAPN+ é essencial para evidenciar a importância do tema na organização. São exemplos de iniciativas:



Organizar workshops interativos com estudos de caso e simulações para que os colaboradores possam compreender os riscos para a população LGBTQIAPN+ e aplicar medidas práticas de governança;



Convidar especialistas em direitos LGBTQIAPN+ e tecnologia para conduzir sessões de conscientização e compartilhar experiências;



Produzir ferramentas e guias práticos (como checklists e questionários) a serem utilizados pelas equipes que atuam no desenvolvimento de sistemas de IA, com o objetivo de mitigar riscos à população LGBTQIAPN+.

MECANISMOS DE MONITORAMENTO E FEEDBACK

A implementação de mecanismos de monitoramento permite que a organização avalie continuamente a eficácia das medidas mitigatórias implementadas, de modo a verificar se estão tendo o impacto desejado na redução dos riscos identificados. Além disso, o monitoramento contínuo permite que a organização identifique e avalie eventuais novos riscos decorrentes de mudanças nos ambientes e condições que envolvem o sistema de IA.



Isso pode ser relevante para coletar feedbacks junto à população LGBTQIAPN+ e identificar oportunidades de melhoria, o que pode ser feito, por exemplo, por meio da criação de canal de comunicação onde pessoas LGBTQIAPN+ possam reportar experiências negativas e oferecer sugestões para aprimoramento da interação com sistemas de IA.

DICA EXTRA:

Para conferir boas práticas de governança em proteção de dados, confira nossa cartilha “Privacidade com Orgulho”, que traz dicas para que ações de diversidade e inclusão continuem a promover os direitos de pessoas LGBTQIAPN+ de forma conjunta com a privacidade.



[LEIA AQUI](#)

É fundamental reconhecer que o desenvolvimento e a utilização de sistemas de IA podem ter impactos significativos na vida das pessoas LGBTQIAPN+. Por meio desta cartilha, procuramos apresentar uma singela contribuição ao tema e destacar a importância de um olhar centrado na comunidade, promovendo a inovação responsável e a construção de um futuro mais inclusivo e equitativo.

