INSTITUTO IGARAPÉ
a think and do tank

MAIO 2022

GUIA PARA USO RESPONSÁVEL, TRANSPARENTE E SEGURO

DA INTELIGÊNCIA ARTIFICIAL NA SEGURANÇA PÚBLICA

Sumário

Introdução 1

Metodologia 3

Definições dos conceitos 4

Princípios e perguntas de orientação 4

Considerações finais 15

Referências 16

Anexo I - Lista de guias, estratégias, projetos de lei, legislações e outros documentos analisados 17

GUIA PARA USO RESPONSÁVEL, TRANSPARENTE E SEGURO
DA INTELIGÊNCIA ARTIFICIAL NA SEGURANÇA PÚBLICA[1]

Introdução

Instituições de segurança pública ao redor do mundo passaram a explorar, nos últimos anos, uma série de novas tecnologias de Inteligência Artificial (IA), definidas aqui como um conjunto de tecnologias que simulam características humanas, tais como conhecimento, resolução de problemas, percepção, aprendizado e planejamento; e também as tecnologias que operam de modo racional. A partir dessas operações, tais tecnologias são capazes de produzir decisões que incluem predição, recomendações e classificações. As tecnologias de IA operam com base em algoritmos que usam as técnicas de aprendizado por máquinas (*machine learning*), incluindo, por exemplo técnicas de aprendizado profundo (*deep learning*), para gerar seus modelos. Assim, plataformas de policiamento preditivo, sistemas de reconhecimento facial, ferramentas que monitoram redes sociais e programas de vigilância online são alguns dos muitos exemplos que vêm sendo desenvolvidos e empregados por órgãos policiais, seja no âmbito da prevenção de ocorrências ou na investigação criminal.

Não obstante, especialistas nos impactos sociais de tecnologias e organizações da sociedade civil vêm apontando os riscos que o entusiasmo das instituições de segurança pública com as tecnologias de IA apresenta para os direitos humanos e liberdades civis, em especial quando se trata dos direitos de determinados grupos vulneráveis da população.[2] Críticos dessas tecnologias apontam que as empresas fornecedoras de tecnologias de IA costumam ressaltar as vantagens dos seus produtos, sem necessariamente mencionar ou considerar a dimensão dos riscos aos direitos e integridade desses grupos mais vulneráveis. Além disso, raramente divulgam evidências científicas concretas de seus impactos e não são transparentes sobre os mecanismos de funcionamento dos seus algoritmos.[3]

No âmbito da segurança pública, os debates e preocupações a respeito do uso de tecnologias de IA trazem questionamentos importantes. O uso de sistemas de reconhecimento facial vêm sendo criticado por suas altas taxas de erro, que resultam, por exemplo, em falsos positivos quando usados sobretudo para identificar mulheres e pessoas negras[4]. Em consequência disso, e das críticas legítimas levantadas por diferentes movimentos por igualdade racial e redução da violência policial, o uso do reconhecimento facial por órgãos de polícia é objeto de um movimento de dimensão global que pede por moratória em seu uso ou seu banimento.[5]

---

1 O Instituto Igarapé agradece aos parceiros que contribuíram com comentários e sugestões em edições preliminares desse documento, em especial o Instituto de Tecnologia e Sociedade (ITS Rio) e a Transparência Brasil.

2 Linder, 2020.

3 Farley, 2020.

4 Nunes, 2021.

5 Access Now, 2021.

Existem ainda as preocupações com relação ao uso de plataformas de policiamento preditivo, que são ferramentas de IA capazes de prever ocorrências criminais a partir de estimativas de probabilidade e risco de crimes cometidos em horários e locais específicos. Há também plataformas preditivas - estas ainda mais criticadas - que são focadas na identificação de perfis individuais que seriam mais propensos a cometerem crimes ou violarem as condições de liberdade condicional. Importante destacar que a literatura científica sobre prevenção e redução de violência indica que intervenções dirigidas para *hot spots* (pontos geográficos com alta concentração de incidentes criminais) é capaz de reduzir crimes.[6] Porém, se mal implementadas e mal administradas, essas plataformas podem reproduzir vieses de caráter social, econômico e racial. O bom desempenho dessas tecnologias vai depender da qualidade dos dados, dos protocolos policiais que alimentam seus modelos e das medidas de identificação e mitigação de vieses colocadas em prática no desenvolvimento e na implementação da ferramenta. O modo como os dados são selecionados e estruturados pode, por si só, refletir práticas discriminatórias e excludentes. No caso da governança dos protocolos policiais, é preciso considerar que a diversidade racial e social das pessoas responsáveis por sua criação, implementação e observância também vai impactar na diminuição de vieses discriminatórios. É necessário que os critérios que orientam esses protocolos sejam claros e auditáveis, levando em consideração o racismo e a discriminação estrutural.

O receio com esses impactos adversos das tecnologias de IA vem gerando um imenso debate que já produz resultados. Diversas organizações ao redor do mundo vêm produzindo cartas de princípios e guias para orientar o desenvolvimento e uso de tecnologias de IA[7]. Vários países já apresentaram suas estratégias para lidar com o desenvolvimento e implementação da Inteligência Artificial, e alguns já se adiantaram com a proposição de legislações para regulamentar o tema.[8] O que todas essas iniciativas têm em comum, porém, é o reconhecimento de que as inovações em IA têm potencial para auxiliar o desenvolvimento humano, desde que desenvolvidas e utilizadas com diligência, cuidado e respeito às liberdades fundamentais.

Contudo, ainda há uma lacuna na produção de documentos de orientação para desenvolvimento e uso de tecnologias de IA. Desde 2013, o Instituto Igarapé vem investigando os impactos da introdução de novas tecnologias na segurança pública no contexto latino-americano[9], concentrando-se principalmente na revisão de literatura, análise empírica e produção de diretrizes sobre câmeras corporais[10] e ferramentas de policiamento preditivo[11] em cidades do Brasil e do restante da América Latina. As especificidades dos diferentes países da América Latina apontam a necessidade de que se desenvolvam diretrizes mais amplas para o uso de tecnologias de IA na segurança pública. Países latino-americanos apresentam as maiores taxas de violência do mundo, advindas de um quadro de extrema desigualdade social. Esse mesmo quadro, por sua vez, traz um cenário propício para episódios constantes de violência policial e violações de direitos humanos contra populações racializadas, mulheres e classes sociais vulneráveis. O uso

6 Braga, Papachristos, Hureay, 2012.

7 Ver documentos e hyperlinks listados no Anexo I.

8 Ver documentos e hyperlinks listados no Anexo I.

9 Willis et al., 2013.

10 Muggah et al., 2016.

11 Aguirre et al., 2019.

indiscriminado de tecnologias de IA , ou seja, aquele que não considera a avaliação e o monitoramento dos impactos sociais de seu uso, por parte das instituições de segurança pública pode se tornar mais um instrumento de reprodução e agravamento dessas diferentes formas de violência.[12]

É com esse cenário em mente que o Instituto Igarapé apresenta aqui o seu Guia para o Uso Transparente, Responsável e Seguro da IA na Segurança Pública. O objetivo deste documento é orientar as Instituições de Segurança Pública para o uso responsável, transparente e seguro de tecnologias de Inteligência Artificial. Entendemos que essas tecnologias podem ser efetivas na promoção da segurança pública, mas é preciso considerar se são adequadas e proporcionais, tendo em vista que podem afetar não somente indivíduos em suas liberdades e dignidade humana, mas também grupos inteiros em sua coletividade.

Esperamos, assim, que este guia sirva também para que gestores e operadores da segurança pública analisem os parâmetros que aqui elencamos e percebam que o estado da arte de determinadas tecnologias dificilmente permitirá que elas possam ser enquadradas nos princípios aqui dispostos. Esse olhar crítico sobre o uso de novas tecnologias vai permitir que elas sejam avaliadas não apenas como um meio de aprimoramento de eficácia e eficiência, mas também como uma oportunidade para refletir a respeito do importante papel que a segurança pública exerce em um Estado Democrático de Direito.

Metodologia

Este guia apresenta seis princípios que devem nortear a aplicação de tecnologias de IA na segurança pública: Transparência, Responsabilidade, Legalidade, Segurança, Privacidade e Equidade (englobando aqui a não-discriminação, a inclusão e a diversidade). Os princípios de Transparência e Responsabilidade, por sua vez, se dividem em parâmetros específicos. Dentro de Transparência, temos os parâmetros da Explicabilidade e da Auditabilidade, enquanto que, dentro do princípio da Responsabilidade, se encontram os parâmetros da Responsividade e da Acurácia.

A definição desses seis princípios foi construída a partir da leitura e análise de 48 documentos, entre guias, recomendações, estratégias nacionais e projetos de lei (Anexo I). A análise de documentos de diversos países nos permitiu compreender quais são os parâmetros que vêm sendo apresentados para a governança da Inteligência Artificial como um todo. Identificamos os princípios que foram propostos e buscamos condensá-los e selecioná-los para chegar em seis princípios pertinentes para as instituições de segurança pública.

Os 48 documentos também nos permitiram elaborar as perguntas que as instituições de segurança pública devem responder no momento em que pretendem adquirir e aplicar uma tecnologia de IA. É a partir dessas perguntas que será possível avaliar se os seis princípios estão sendo atendidos durante a utilização das ferramentas adquiridas.

12 Benjamin, 2019; Heaven, 2020.

Definições dos conceitos

Os conceitos que apresentamos abaixo são centrais para a compreensão e aplicação deste guia. Contudo, é importante ressaltar que são conceitos amplos, passíveis de serem definidos de muitas formas. Assim, sem nenhum prejuízo aos debates em torno desses conceitos, as definições que propomos aqui servem apenas para orientar a leitura e entendimento deste documento.

Inteligência Artificial (IA): o termo se refere a um conjunto de tecnologias que simulam características humanas, tais como conhecimento, resolução de problemas, percepção, aprendizado e planejamento; como também se refere a tecnologias que operam de modo racional. São capazes de, a partir dessas operações, produzirem decisões que incluem predição, recomendações e classificações. Tecnologias de IA operam com base em algoritmos que usam as técnicas de aprendizado por máquinas (*machine learning*) ou de aprendizado profundo (*deep learning*) para gerar seus modelos.

Fornecedores de tecnologias de IA: são instituições, de caráter público - podendo incluir as próprias Instituições de Segurança Pública - ou privado, que desenvolvem e/ou fornecem sistemas de soluções ou aplicações tecnológicas que fazem uso de IA.

Instituições de segurança pública: autoridade pública, órgão ou entidade do Poder Público - seja ele federal, estadual ou municipal - responsável pela prevenção, detecção, investigação ou repressão de atos infracionais e infrações penais ou execução de sanções penais, incluindo a salvaguarda e a prevenção de ameaças à segurança pública.

Atividade de segurança pública: toda e qualquer atividade exercida para a preservação da ordem pública, usufruto de direitos e para a prevenção e detecção de infrações penais, realizada por autoridades legalmente competentes.

Indivíduos: pessoa natural, ou grupo de pessoas naturais, que serão diretamente afetadas pela aplicação de tecnologias de IA.

Aplicação: trata-se da utilização de tecnologias de IA por parte das Instituições de segurança pública, providas por fornecedores de tecnologias de IA, com um ou mais fins específicos e previamente determinados.

Princípios e perguntas de orientação

Transparência

O conceito de transparência é extremamente amplo. De modo geral, podemos afirmar que ele diz respeito às qualidades de abertura e acessibilidade de determinadas práticas, políticas e conhecimentos. Quando falamos de abertura em tecnologias de IA empregadas na segurança pública, nos referimos aos processos de concepção, desenvolvimento, design, aplicação e monitoramento. Assim, a abertura implica em dois esforços distintos, direcionados a diferentes públicos: de um lado, a explicabilidade desses sistemas para um público amplo e leigo, principalmente aqueles que serão impactados pela ferramenta, e, do outro lado, a capacidade de serem auditados de modo independente por especialistas. Finalmente, as instituições de segurança pública devem priorizar o emprego de tecnologias de IA de código aberto ou exigir de seus fornecedores de solução uma

documentação adequada e acessível, que permita a qualquer pessoa com conhecimento técnico verificar o funcionamento dos seus algoritmos, facilitando assim o entendimento sobre a aplicação. Da mesma forma, no caso de aplicações de ferramentas privadas, os gestores devem exigir dos seus fornecedores explicações transparentes sobre o funcionamento dos algoritmos e que os mesmos possam ser auditados.

a. Explicabilidade

Um importante aspecto na aplicação de tecnologias de IA para fins de segurança pública é a explicação dos processos de decisão imbuídos nesses sistemas. Os indivíduos afetados possuem o direito de compreender as decisões informadas direta ou indiretamente por IA, bem como quais são os dados que orientam essas decisões. Quando possível, os responsáveis pela aplicação devem realizar esforços para explicar esses aspectos de modo não-técnico e sem ambiguidades. É importante lembrar que, explicar o processo de decisão de uma IA não implica na abertura do seu código-fonte ou outros segredos de negócio. Na realidade, revelar o código-fonte não necessariamente é suficiente para explicar as regras que orientam as decisões de uma IA.

Em muitos casos, é mais benéfico aos indivíduos conhecerem os processos que informam essas decisões. Da mesma forma, é importante deixar claro quem são os responsáveis pelo desenvolvimento da aplicação, quem são os responsáveis por operá-la, quem são os responsáveis por sua manutenção e quais serão os indivíduos afetados. Quando os responsáveis pela aplicação anunciam um novo serviço ou funcionalidade gerados por um processo de aprendizado de máquinas, reduzir a explicação para um simples termo como “aprendizado de máquinas” ou “sistema de decisão automática” impede que os indivíduos sejam capazes de distinguir entre efeitos aleatórios e efeitos mais perniciosos da decisão de uma IA.

Instituições de segurança pública que pretendem adquirir e empregar tecnologias de IA devem se orientar pelas seguintes questões, visando atender aos critérios de explicabilidade:

Descrição dos algoritmos envolvidos na aplicação?
Sempre que possível, os algoritmos que sustentam a aplicação devem ser explicados publicamente para toda a população. Caso sejam tecnologias de código aberto, este deve ser disponibilizado. Porém, mais importante do que isso, o funcionamento do algoritmo deve ser explicado de modo não-técnico e de fácil compreensão.

A comunidade afetada foi devidamente consultada sobre a aplicação?
Os indivíduos que serão diretamente afetados pela aplicação precisam ser notificados ou, preferencialmente, consultados de forma pública e transparente ao longo do processo de implementação da ferramenta. Isso pode ser feito nos sítios eletrônicos das instituições de segurança pública e através de veículos de mídia, mas é importante que a comunicação seja ativa e também ocorra nos locais de aplicação da ferramenta. Da mesma forma, após o início da aplicação, quaisquer novos desenvolvimentos, funcionalidades ou evoluções em projetos piloto também precisam ser comunicadas.

Quais são os dados utilizados na aplicação?
Além das explicações sobre o funcionamento dos algoritmos, as instituições de segurança pública devem informar os indivíduos sobre quais dados são utilizados para alimentar a aplicação de IA, seja no seu treinamento, no próprio emprego da tecnologia e mesmo quais dados serão armazenados. A explicação sobre os dados, incluindo aí explicações sobre como são utilizados nos processos de tomada

de decisão, torna-se ainda mais imperativa quando se tratam de dados pessoais.

Quais são os indicadores de impacto da aplicação?
Indicadores de impacto são referências quantitativas e qualitativas que nos permitem avaliar os efeitos e mudanças gerados pela aplicação na realidade. Tradicionalmente, fornecedores de tecnologia para a segurança pública apresentam indicadores de impacto operacional, ou seja, aqueles que demonstram a eficácia e a eficiência das ferramentas. Contudo, é importante que os fornecedores também apresentem às instituições de segurança pública os indicadores de impacto social de suas tecnologias. Por impacto social, entendemos aqui as possíveis transformações positivas e negativas que podem afetar os grupos no contexto da aplicação. Assim, os indicadores de impacto social tornam possível identificar e monitorar os impactos da aplicação de IA. Da mesma forma, sempre que possível, esses indicadores devem ser comunicados ao público, bem como os protocolos de monitoramento dos impactos e protocolos de governança da aplicação.

Quais serão os atores e grupos afetados pela aplicação?
Aplicações de IA na segurança pública sempre terão pessoas ou grupos afetados direta ou indiretamente. Cada aplicação acontecerá em locais e contextos diversos. As instituições responsáveis precisam ter o mapeamento desses locais e indivíduos, para consultá-los e informá-los de forma clara e precisa os motivos do resultado da ferramenta, antes de iniciarem a aplicação. Assim, poderão informá-los sobre os objetivos, resultados esperados, formas de monitoramento e de compartilhamento dos resultados. Também é importante que as instituições de segurança pública se mostrem abertas para ouvir grupos impactados, mas que não constaram no mapeamento inicial.

Quais são os objetivos da aplicação?
Toda aplicação de IA tem uma finalidade e objetivos a serem alcançados. Estes devem ser comunicados ao público e, se porventura vierem a ser alterados posteriormente, tais transformações também devem ser comunicadas.

b. Auditabilidade

Um aspecto específico da explicabilidade de uma aplicação de IA é sua auditabilidade. Aplicações devem ser submetidas à avaliação e revisão dos seus processos de decisão por parte de indivíduos e grupos independentes. Isso exige que informações relevantes sejam disponibilizadas para permitir a testagem, monitoramento e produção de pareceres. Da mesma forma, é preciso encorajar processos de auditoria independente e periódica durante toda a vida da aplicação. O objetivo é informar os indivíduos que a aplicação da IA foi auditada por um agente independente e continua aberta à novas auditorias no futuro, levando em consideração que essas ferramentas precisam ser calibradas e revistas constantemente para manterem sua acurácia.

Para atender aos critérios do princípio da auditabilidade, instituições de segurança pública que pretendem adquirir e empregar tecnologias de IA devem se orientar pelas seguintes questões:

Houve alguma forma de auditoria independente do sistema aplicado?
No momento de adquirir e empregar uma tecnologia de IA, instituições de segurança pública devem priorizar sistemas que foram submetidos a auditorias independentes. Os fornecedores precisam fornecer os resultados dessas auditorias e, quando possível, estes também devem ser comunicados ao público.

Há a possibilidade novas auditorias públicas ou feitas por terceiros independentes?
Para além das primeiras auditorias, uma aplicação transparente deve estar aberta para novas auditorias. Estas podem ser feitas por organizações independentes contratadas pelas instituições de segurança pública, mas também podem ser realizadas a pedidos da sociedade civil. Inclusive, a realização de auditorias constantes, com intervalos regulares, é considerada uma boa prática.

Responsabilidade

A responsabilidade[13] sobre as aplicações de IA na segurança pública diz respeito aos funcionários, supervisores e equipes responsáveis, de um lado, pela concepção, desenvolvimento e *design* das tecnologias, mas também àqueles que irão operar os sistemas quando forem utilizados. Ou seja, a responsabilidade pelas formas de aquisição, uso e qualidade dos dados é das instituições de segurança pública. Os procedimentos que orientam as ações dessas pessoas terão uma influência direta nos impactos das aplicações de IA. No caso das aplicações na segurança pública, entendemos que a responsabilidade se refere principalmente aos seguintes aspectos: a capacidade dos indivíduos afetados contactarem as instituições de segurança pública responsáveis pela operação da aplicação, incluindo o conhecimento sobre como fazê-lo e a garantia de que terão suas demandas respondidas; e o comprometimento dos responsáveis pelo fornecimento da tecnologia com a acurácia das aplicações e com seus impactos sociais, sejam eles intencionais ou não.

Para garantir que os critérios de responsabilidade serão cumpridos durante uma aplicação de IA, as instituições de segurança pública devem buscar as respostas para as seguintes questões:

Quais são os benefícios, desafios e riscos envolvidos?
Antes de iniciar qualquer aplicação de IA, instituições de segurança pública devem fazer um mapeamento extensivo dos benefícios, desafios a serem enfrentados e riscos envolvidos no processo. Os riscos devem ser entendidos de maneira ampla, incluindo os riscos de exclusão e impactos negativos em direitos humanos. Este mapeamento é necessário para determinar a cadeia de responsabilidade sobre a aplicação e deve ser feito desde o momento da aquisição da tecnologia, e continuar durante as fases de implementação, uso e monitoramento de resultados. Quando possível, benefícios, desafios e riscos devem ser comunicados ao público.

Há algum corpo na estrutura organizacional responsável pela supervisão da aplicação?
Aplicações de IA na segurança pública tendem a apresentar muitos riscos, tendo em vista que tratam de temas sensíveis e têm potencial para amplificar a violação de direitos fundamentais. Em consequência disso, é necessário que as instituições responsáveis criem um órgão de supervisão das aplicações de IA, dentro de sua estrutura organizacional. É altamente recomendável que, quando houver a possibilidade, os órgãos de supervisão incluam representantes das comunidades afetadas e de outros especialistas. Os representantes comunitários e especialistas darão voz às preocupações dos grupos afetados e podem ser uma garantia de que os impactos sociais sejam considerados desde a aquisição da tecnologia.

13 A responsabilidade que aqui se refere está relacionada ao dever de responder por determinada ação, não estando associada ao conceito jurídico de responsabilidade civil, ou seja, não diz respeito à obrigação legal de reparar danos causados a terceiros prevista no Código Civil Brasileiro.

Houve algum treinamento ou preparo da pessoa (ou pessoas) responsável (is)?
Os responsáveis pela aplicação de IA na segurança pública devem passar treinamento voltado à tecnologia empregada. Além disso, é importante que esse treinamento inclua o conhecimento sobre os desafios e riscos envolvidos na aplicação, para que estejam cientes e atentos aos problemas que venham a ocorrer.

Há supervisão ativa de um agente humano no processo de avaliação e decisão da IA utilizada na aplicação?
A autonomia dos agentes humanos no curso de uma aplicação de IA é essencial, principalmente quando se trata do emprego dessa tecnologia na segurança pública. Dessa forma, é possível atribuir responsabilidades claras sobre eventuais erros e impactos adversos da aplicação. Agentes humanos devem supervisionar o desempenho de uma IA e ser os responsáveis finais sobre quaisquer decisões tomadas com auxílio de tecnologias. Além disso, mesmo nos casos em que há um agente humano no curso da aplicação, é preciso garantir que a supervisão tenha capacidade para rejeitar e influenciar o resultado das decisões fornecidas pela IA. Trata-se de uma supervisão ativa do processo decisório que aplicará esses resultados. É importante que os fornecedores da aplicação deem apoio no desenho, monitoramento e eventuais alterações deste processo.

Existe alguma determinação clara sobre a cadeia de responsabilidade, caso a aplicação venha a lesionar os direitos fundamentais de um indivíduo?
As instituições de segurança pública, desde o momento em que concebem ou iniciam a aplicação de uma tecnologia de IA, precisam determinar a cadeia de responsabilidade sobre as decisões que serão tomadas no seu emprego. A cadeia de responsabilidade pode incluir representantes dos fornecedores da tecnologia usada na aplicação.

a. Responsividade

Indivíduos devem ter a possibilidade de contactar os responsáveis pela aplicação dentro das instituições de segurança pública para manifestar suas dúvidas, questionamentos e pedir explicações sobre preocupações legítimas concernentes às decisões de uma IA, bem como sobre os protocolos de uso e tomada de decisão com base nos resultados dessa aplicação. Instituições de segurança pública devem providenciar meios claros e acessíveis para que indivíduos aprendam mais sobre essas ferramentas. Além disso, esses meios precisam estar disponíveis para os casos em que indivíduos apontem um efeito adverso, resultado de processos decisórios baseados na aplicação de uma IA.

Para garantir que as aplicações de IA atenderão aos requisitos básicos de responsividade, as instituições de segurança pública devem responder às seguintes questões:

Quem é o responsável por atender o indivíduo, caso este se sinta lesado durante a aplicação?
É fundamental que as instituições de segurança pública designem uma ou mais pessoas responsáveis para atendimento específico do público que venha a ter dúvidas ou se sinta lesado por uma aplicação de IA. É importante definir e divulgar como será o processo de atendimento, incluindo os canais de atendimento e resposta, o tempo de retorno e eventuais possibilidades de recurso.

Como o indivíduo pode contactar o responsável no caso de se sentir lesado?
O contato com o pessoal responsável por atender o público, dirimir dúvidas, endereçar reclamações e pedidos de reavaliações de decisões e correções deve ser de fácil acesso. Além disso, esse canal deve ser amplamente divulgado pelas instituições de segurança pública.

b. Acurácia

Indivíduos devem ter a possibilidade de conhecer a acurácia de uma aplicação de IA. É sabido que a acurácia é uma preocupação entre desenvolvedores de tecnologias de IA, dada a complexidade dos algoritmos e as incertezas envolvidas nos processos de aplicação. Contudo, é recomendável que os responsáveis pela aplicação e fornecedores sejam transparentes sobre o modo como identificam, registram e reportam quaisquer fontes de erro, bem como com relação ao tratamento de dados empregado para mitigação de impactos sociais negativos e os dados que foram removidos para redução de vieses na aplicação. Recomenda-se também que criem meios para disseminar os resultados dos testes de acurácia e os processos de mitigação de quaisquer erros encontrados.

Instituições de segurança pública que pretendem adquirir e empregar tecnologias de IA devem avaliar o princípio da acurácia orientadas pelas seguintes questões:

Existem dados sobre a acurácia da aplicação?
As instituições de segurança pública devem sempre exigir dos fornecedores de tecnologia dados sobre a acurácia da aplicação que será adquirida, inclusive com dados específicos para diferentes grupos populacionais. Também é importante que os dados de acurácia apresentados pelo fornecedor sejam contextualizados com a acurácia esperada para o tipo de tecnologia em questão, comparando-os com soluções semelhantes.

Os dados sobre a avaliação de acurácia da aplicação estão disponíveis?
A avaliação da acurácia da aplicação precisa ser disponibilizada ao público, para que este tenha ciência sobre a precisão da tecnologia de IA que será empregada. Ademais, as avaliações e monitoramento da acurácia devem ser realizados constantemente, mesmo após o início da aplicação, como modo de garantir que os níveis de acurácia são satisfatórios.

Legalidade

Toda aplicação de tecnologias de IA na segurança pública devem estar embasadas em hipótese legal. Os termos desse embasamento podem variar de acordo com a aplicação e a instituição de segurança pública responsável, mas é fundamental que estas tenham um respaldo legal para fazê-lo. As instituições de segurança pública devem ter uma compreensão das leis que regulamentam suas atividades e, por consequência, do modo como estas se aplicam na utilização de tecnologias de IA em particular. Assim, a mera disponibilidade e acesso à tecnologia não podem servir como justificativa para sua utilização. O respeito ao princípio da legalidade pode ser verificado pelas seguintes questões:

Qual o embasamento legal da aplicação?
As instituições de segurança pública têm plena consciência dos seus deveres e limites determinados por lei. No entanto, é preciso que essas considerações sejam observadas e reavaliadas no curso de qualquer aplicação de tecnologias de IA. Isso significa que essas ferramentas podem ser utilizadas quando se mostrarem necessárias para o cumprimento das atribuições legais das instituições de segurança pública, tendo sempre o interesse público como objetivo final. Órgãos de investigação, por exemplo, possuem atribuição legal que justifica o uso de ferramentas de IA que auxiliem no monitoramento, desde que preservem os direitos individuais e que não existam outros meios eficientes para tal. Essa avaliação, que identifica os direitos individuais afetados e a existência de outros meios menos arriscados e proporcionais para o cumprimento das atribuições das instituições de segurança pública, deve ser a principal baliza que fundamenta a decisão de adquirir e aplicar uma ferramenta

de IA. Assim, é recomendado que exista um processo de comunicação interno e, principalmente, externo que explique como determinada aplicação é necessária, adequada e proporcional para que a instituição de segurança cumpra com suas obrigações.

Há necessidade de autorização judicial para garantir a legalidade da aplicação?
Algumas aplicações de tecnologias de IA podem exigir autorização judicial prévia. Casos que envolvam a suspensão de direitos individuais, como o monitoramento de indivíduos, por exemplo, necessitam de autorização judicial para garantir as salvaguardas legais, tendo em vista que essa atividade por parte do Estado é extremamente intrusiva e potencializadora de violações.

Segurança

Aplicações de tecnologias de IA na segurança pública devem adotar sistemas desenvolvidos tendo a segurança digital como prioridade desde o momento de sua criação, ou seja, seguros por *design (secure by design)*. Além disso, as aplicações devem garantir a segurança digital dos ativos que os compõem e dos indivíduos afetados. Finalmente, para garantir a segurança digital da aplicação, é preciso que existam controles de acesso dentro do quadro de funcionários das instituições de segurança pública e das empresas responsáveis pela aplicação. Isso implica, como ponto de partida, na criação de listas, níveis e registros de acesso, bem como na justificativa dos motivos pelos quais esses funcionários acessam a tecnologia. Os registros de acesso devem ser armazenados por tempo determinado.

Para verificar se uma aplicação de IA atende aos critérios de segurança digital, instituições de segurança pública que pretendem adquirir e empregar essas tecnologias de IA devem se orientar pelas seguintes questões:

Existem protocolos claros para a transferência de dados entre os usuários da aplicação?
É comum que aplicações de IA envolvam a transferência de dados entre seus usuários. Nos casos em que essa aplicação ocorre para fins de segurança pública, existe a transferência de bases de dados entre diferentes órgãos públicos e entre o órgão responsável e o fornecedor da tecnologia. Todas as transferências devem apresentar protocolos claros e que atendam aos melhores padrões de segurança da informação ao alcance dos atores envolvidos.

Existe menção explícita aos padrões de segurança da informação empregados no curso da aplicação?
Os indivíduos afetados precisam ser comunicados sobre os padrões de segurança da informação empregados na aplicação de IA. Evidentemente, o modo como essa comunicação será feita deve levar em consideração a sensibilidade e o segredo envolvidos em algumas atividades de segurança. Ainda assim, dado os riscos envolvidos e as potenciais vulnerabilidades dos sistemas de IA, o público precisa ter a garantia de que seus dados estão protegidos.

Privacidade

A proteção de dados pessoais e o direito à privacidade se tornaram uma preocupação legítima em todo o mundo. Proteger as informações dos indivíduos afetados por aplicações de IA, especialmente na segurança pública, vem se tornando obrigatória em muitas jurisdições, incluindo a necessidade de se explicar a política de tratamento e retenção dos dados adotada. Sempre que possível, é fundamental comunicar aos indivíduos a respeito da segurança dos seus dados e a resiliência dos sistemas. Todas as aplicações de IA utilizadas na segurança pública devem ser desenvolvidas tendo a privacidade como elemento essencial, bem como devem explicar como as tecnologias adotadas estão de acordo com as leis de proteção de dados vigentes.

Instituições de segurança pública que pretendem adquirir e empregar tecnologias de IA devem se orientar pelas seguintes questões, visando atender aos critérios de privacidade:

Como o sistema aplicado atende à legislação de proteção de dados vigente?
As instituições de segurança pública devem exigir que os fornecedores de tecnologias de IA demonstrem como o produto oferecido atende à legislação de proteção de dados vigente nos locais da aplicação. Da mesma forma, as instituições devem garantir ao público, de forma clara e didática, que a aplicação em curso respeita os direitos fundamentais à privacidade e à proteção de dados. Também é preciso criar protocolos de acesso, armazenamento e compartilhamento dessas informações.

Existem medidas para anonimização de dados pessoais, caso estes sejam utilizados?
A utilização de dados pessoais para fins de segurança pública, principalmente quando trata de dados pessoais sensíveis, deve ser feita de modo a garantir sua anonimização. Assim, é possível utilizar a tecnologia para facilitar o trabalho das instituições de segurança e, ao mesmo tempo, preservar os direitos fundamentais à privacidade e proteção de dados.

Equidade

A equidade vem se tornando o aspecto mais relevante quando se fala em processos de decisão tomados por meio de aplicações de IA. Apesar de existirem muitas definições sobre o que seria essa "equidade", adotamos aqui uma definição mais ampla. A equidade deve incluir a não discriminação, a inclusão e a diversidade. Decisões de IA que prezam pela equidade não devem gerar impactos injustos e discriminatórios (sejam eles com relação à raça, gênero, condições físicas e mentais, idade ou classe social). Ademais, os indivíduos afetados e o público em geral devem ser informados sobre as medidas de mitigação de vieses utilizadas pelos fornecedores e instituições de segurança pública envolvidos na aplicação, evitando assim decisões das IA que sejam discriminatórias e prejudiciais a grupos vulneráveis da população.

Para garantir a equidade das aplicações de IA, as instituições envolvidas devem consultar com especialistas locais e representantes da população que pode vir a ser potencialmente afetada por impactos negativos de uma aplicação. Esses especialistas e representantes conhecem a comunidade local, suas dinâmicas e o modo como podem ser impactadas. Ouvir as considerações desses grupos e acolher suas preocupações pode trazer maior legitimidade à aplicação. Finalmente, um aspecto específico para garantir a equidade nas aplicações de IA diz respeito ao desenvolvimento e treinamento ético das forças de segurança que irão empregar essas tecnologias.

Instituições de segurança pública que pretendem adquirir e empregar tecnologias de IA devem se orientar pelas seguintes questões, visando atender aos critérios de equidade:

Existem grupos sociais que podem ser prejudicados ou privilegiados no contexto de aplicação do sistema?
No processo de avaliação dos desafios e riscos envolvidos na aplicação de tecnologias de IA, as instituições de segurança pública precisam levar em consideração diferentes recortes populacionais. A análise de risco deve ponderar como os impactos da aplicação irão afetar grupos vulneráveis de modo distinto, permitindo inclusive que especialistas sejam ouvidos nessa avaliação. É importante garantir que os impactos benéficos sejam inclusivos. Análises desse tipo devem ser realizadas a partir de testes técnicos, tais como os relatórios de impacto, que mencionaremos adiante.

Alguma medida foi empregada para a mitigação desses efeitos?
A avaliação dos impactos sociais deve ser acompanhada de medidas de mitigação dos efeitos negativos, em particular daqueles que afetam grupos vulneráveis. Essas medidas precisam ser realizadas em parceria com instituições de segurança e fornecedores. Caso identifique-se que a mitigação dos impactos sociais negativos não é satisfatória, a aplicação deve ser suspensa.

Quais são as bases de dados utilizadas? Elas são diversas, representativas e apropriadas para o contexto da aplicação? Há algum protocolo quanto a isso?
Instituições de segurança pública e fornecedores de tecnologia de IA devem garantir que as bases de dados utilizadas no treinamento e no emprego da aplicação sejam suficientemente diversas em termos raciais, de gênero e de classe social, evitando assim a geração de vieses perniciosos para grupos vulneráveis.

Houve a apresentação de relatório de impacto?
Relatórios de impacto são voltados para a identificação, prevenção e mitigação de impactos negativos das aplicações de tecnologias de IA. Instituições de segurança pública devem priorizar fornecedores de tecnologia de IA que apresentem relatórios de impacto sobre seus produtos. Da mesma forma, é recomendável que as instituições realizem relatórios de impacto sobre suas aplicações periodicamente.

Existem protocolos que garantem a qualidade dos dados, evitando assim a criação de vieses nos algoritmos da IA?
É recomendado que as instituições de segurança pública desenvolvam protocolos que garantam a qualidade dos dados que irão alimentar os algoritmos da aplicação, seja em seu treinamento, na fase de emprego e durante o monitoramento da aplicação. A qualidade dos dados precisa garantir que estes sejam inclusivos e diversos, de modo a evitar a produção de resultados enviesados. Os protocolos também podem ser desenvolvidos em parceria com os fornecedores das tecnologias, principalmente na fase de emprego da aplicação, tendo em vista que muitas vezes a qualidade dos dados vai depender da atuação dos agentes de segurança.

Há diversidade nas equipes de concepção, design, desenvolvimento e aplicação?
Existe um consenso de que inclusão e diversidade social no desenvolvimento de tecnologias de IA é um dos elementos que pode garantir a criação de sistemas éticos e que respeitam os direitos humanos. A diversidade de olhares raciais, de gênero e sociais permite que problemas de discriminação sejam identificados no momento de concepção de uma tecnologia. Assim, as instituições de segurança pública devem priorizar tecnologias de fornecedores que apresentem essa diversidade em suas equipes de desenvolvimento. Da mesma forma, as equipes envolvidas na aplicação da tecnologia dentro das instituições de segurança também devem levar em consideração essa diversidade em sua formação.

Princípios		Perguntas de orientação
Transparência	Explicabilidade	Descrição dos algoritmos envolvidos na aplicação
		A comunidade afetada foi devidamente consultada sobre a aplicação?
		Quais são os dados utilizados na aplicação?
		Quais são os indicadores de impacto da aplicação?
		Quais serão os atores e grupos afetados pela aplicação?
		Quais são os objetivos da aplicação?
	Auditabilidade	Houve alguma forma de auditoria independente do sistema aplicado?
		Há a possibilidade novas auditorias públicas ou feitas por terceiros independentes?
Responsabilidade	Aspectos gerais	Quais são os benefícios, desafios e riscos envolvidos?
		Há algum corpo na estrutura organizacional responsável pela supervisão da aplicação?
		Houve algum treinamento ou preparo da pessoa (ou pessoas) responsáveis?
		Há supervisão ativa de um agente humano no processo de avaliação e decisão da IA utilizada na aplicação?
		Existe alguma determinação clara sobre a cadeia de responsabilidade, caso a aplicação venha a lesionar os direitos fundamentais de um indivíduo?
	Responsividade	Quem é o responsável por atender o indivíduo, caso este se sinta lesado durante a aplicação?
		Como o indivíduo pode contactar o responsável no caso de se sentir lesado?
	Acurácia	Existem dados sobre a acurácia da aplicação?
		Os dados sobre a avaliação de acurácia da aplicação estão disponíveis?
Legalidade		Qual o embasamento legal da aplicação?
		Há necessidade de autorização judicial para garantir a legalidade da aplicação?
Segurança		Existem protocolos claros para a transferência de dados entre os usuários da aplicação?
		Existe menção explícita aos padrões de segurança da informação empregados no curso da aplicação?

continuação

Princípios		Perguntas de orientação
Privacidade		Como o sistema aplicado atende à legislação de proteção de dados vigente?
		Existem medidas para anonimização de dados pessoais, caso estes sejam utilizados?
Equidade		Existem grupos sociais que podem ser prejudicados ou privilegiados no contexto de aplicação do sistema?
		Alguma medida foi empregada para a mitigação desses efeitos?
		Quais são as bases de dados utilizadas? Elas são diversas, representativas e apropriadas para o contexto da aplicação? Há algum protocolo quanto a isso?
		Houve a apresentação de relatório de impacto?
		Existem protocolos que garantem a qualidade dos dados, evitando assim a criação de vieses nos algoritmos da IA?
		Há diversidade nas equipes de concepção, design, desenvolvimento e aplicação?

Considerações finais

Os princípios e orientações que apresentamos neste guia são fruto da análise de uma série de documentos nacionais e internacionais. Estes representam um extenso e intenso debate que busca ponderar as preocupações de grupos vulneráveis e da sociedade civil, com as necessidades das instituições de segurança pública e as transformações tecnológicas.

Os debates fazem com que o campo da ética e transparência na Inteligência Artificial continue sendo um espaço diverso e em transformação. Assim, nosso documento não busca encerrar essa discussão, consolidando princípios que pensamos ser definitivos. Mais do que isso, o objetivo foi trazer o debate para as instituições de segurança pública e situar seus agentes nessa discussão.

Ainda há espaço para aperfeiçoamento e ajustes nas definições que apresentamos e no modo de concretizá-las. Esperamos assim que a aplicação deste guia sirva como mais um elemento na construção de uma segurança pública focada na prevenção da violência e conectada às discussões sobre novas tecnologias e seus impactos na preservação dos direitos humanos.

Referências

Access Now (2021). *Open letter calling for a global ban on biometric recognition technologies that enable mass and discriminatory surveillance.* Disponível em: https://www.accessnow.org/cms/assets/uploads/2021/06/BanBS-Statement-English.pdf

Aguirre, K., Badran, E., Muggah, R. (2019). *Future Crime: Assessing twenty first century crime prediction.* Disponível em: https://igarape.org.br/wp-content/uploads/2019/07/2019-07-12-NE_33_Future_Crime.pdf

Benjamin, R. (2019). Assessing Risk, Automating Racism. *Science*, v. 366, n. 6464.

Braga, A., Papachristos, A., Hureau, D. (2012). *Hot Spots Policing Effects on Crime*. Disponível em: http://www.campbellcollaboration.org/media/k2/attachments/Braga_Hot_Spots_Policing_Review.pdf

Farley, A. (2020). Meet the computer scientist and activist who got Big Tech to stand down. *Fast Company*. Disponível em: https://www.fastcompany.com/90525023/most-creative-people-2020-joy-buolamwini

Heaven, W. D. (2020). Predictive policing algorithms are racist. They need to be dismantled. *MIT Technology Review*. Disponível em: https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/?fbclid=IwAR3zTH9U0OrjaPPqifYSjldzgqylbag6m-GYKBAPQ7jo488SYYl5NbfzrjI

Linder, C. (2020). Why hundreds of mathematicians are boycotting predictive policing. *Popular Mechanics*. Disponível em: https://www.popularmechanics.com/science/math/a32957375/mathematicians-boycott-predictive-policing/

Muggah, R., Badran, E., Siqueira, B., Kosslyn, J. (2016). *Filling the accountability gap: principles and practices for implementing body cameras for law enforcement.* Disponível em: https://igarape.org.br/wp-content/uploads/2017/01/AE-23_Filling-the-accountability-gap-body-worn-cameras-05-01.pdf

Nunes, P. (2021). O Algoritmo e Racismo Nosso de Cada Dia. *Piauí*. Disponível em: https://piaui.folha.uol.com.br/o-algoritmo-e-racismo-nosso-de-cada-dia/

Willis, G.D., Muggah, R., Kosslyn, J., Leusin, F. (2013). *Smarter Policing: Tracking the Influence of New Information Technology in Rio de Janeiro.* Disponível em: https://igarape.org.br/wp-content/uploads/2013/10/Smarter_Policing_ing.pdf

Anexo I - Lista de guias, estratégias, projetos de lei, legislações e outros documentos analisados

Brasil	
Estratégia Brasileira de Transformação Digital	https://www.gov.br/mcti/pt-br/centrais-de-conteudo/comunicados-mcti/estrategia-digital-brasileira/estrategiadigital.pdf
PL 5051/2019	https://legis.senado.leg.br/sdleg-getter/documento?dm=8007560&ts=1594036674731&disposition=inline
PL 5691/2019	https://legis.senado.leg.br/sdleg-getter/documento?dm=8030917&ts=1594037339035&disposition=inline
Consultoria para a Estratégia Brasileira de Inteligência Artificial	https://antigo.mctic.gov.br/mctic/export/sites/institucional/inovacao/paginas/politicasDigitais/Inteligencia/Arquivo/Consultoria-IA-Produtos-I-e-II.pdf
Estratégia Brasileira de Inteligência Artificial (EBIA) 2021	https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/transformacaodigital/arquivosinteligenciaartificial/ia_estrategia.pdf
Conselho Nacional de Justiça (CNJ) - Resolução nº 332, de 21 de Agosto de 2020.	https://atos.cnj.jus.br/files/original191707202008255f4563b35f8e8.pdf
PLC 21/2020	https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1853929
PL 240/2020	https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1857144&filename=Tramitacao-PL+240/2020
Diálogos UE-Brasil: IA e Regulação de Algoritmos 2018	http://www.sectordialogues.org/documentos/proyectos/adjuntos/49f7d3_Intelig%C3%AAncia%20Artificial%20e%20Regula%C3%A7%C3%A3o%20de%20Algoritmos.pdf
Estados Unidos	
US Self Drive Act 2017	https://www.congress.gov/bill/115th-congress/house-bill/3388/text
US Algorithmic Accountability Act	https://www.congress.gov/bill/116th-congress/house-bill/2231/text
US Future of AI Act 2020/ “Fundamentally Under standing the Usability and Realistic Evolution of Artificial Intelligence Act of 2020”	https://www.congress.gov/bill/116th-congress/senate-bill/3771/text
Guidelines for Ethical Development of AI 2019	https://www.congress.gov/bill/116th-congress/house-resolution/153/text
US Preparing for the Future of AI 2016	https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf
White House Summit on AI for American Industry 2018	https://www.hsdl.org/?abstract&did=811092

continuação

União Européia	
Liability for emerging digital tech 2018	https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A52018SC0137
Orientações éticas para uma IA de confiança 2019	https://op.europa.eu/en/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1/language-pt/format-PDF
European Parliament recommendations to the Comission on Civil Law Rules on Robotics (2013/2015)	https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html
European Commission - On Artificial Intelligence - A European approach to excellence and trust 2020	https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
Comunicação da Comissão Europeia: IA para a Europa (2018)	https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52018DC0237&from=EN
A European approach to Artificial intelligence: Coordinated Plan on Artificial Intelligence 2021 Review	https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review
Ethics guidelines for trustworthy AI	https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
The ethics of artificial intelligence: Issues and initiatives	https://www.europarl.europa.eu/RegData/etudes/STUD/2020/634452/EPRS_STU(2020)634452_EN.pdf
European Commission in Key requirements for trustworthy AI, Apr 8, 2019	https://ec.europa.eu/transparency/regdoc/rep/1/2019/EN/COM-2019-168-F1-EN-MAIN-PART-1.PDF
Reino Unido	
UK Robotics and AI 5th report 2016-2017	https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/896/896.pdf
Japão	
New Robot Strategy (Japan) 2015	https://www.meti.go.jp/english/press/2015/pdf/0123_01b.pdf
Canadá	
Building an AI World: Report on National	https://cifar.ca/wp-content/uploads/2020/10/building-an-ai-world-second-edition.pdf
China	
China Notice on the Issuance of Next Generation AI Development Plan 2017	https://www.newamerica.org/cybersecurity-initiative/digichina/blog/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/
Beijing AI Principles 2019	https://www.baai.ac.cn/news/beijing-ai-principles-en.html
Singapura	
Singapore - Model AI Governance Framework (2020)	https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf

continuação

França	
France - For a Meaningful AI 2017/2018	https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
Organização das Nações Unidas (ONU)	
UNESCO - First Draft of the Recommendation on the Ethics of Artificial Intelligence 2020	https://unesdoc.unesco.org/ark:/48223/pf0000373434
Resource Guide on Artificial Intelligence (AI) Strategies- Department of Economic and Social Affairs Sustainable Development	https://sdgs.un.org/sites/default/files/2021-04/Resource%20Guide%20on%20AI%20Strategies_April%202021_rev_0.pdf
ITU - Responsible AI: designing AI for Human Values 2017	http://www.itu.int/pub/S-JOURNAL-ICTS.V1I1-2017-1
Organização para a Cooperação e Desenvolvimento Econômico (OCDE) OECD Principles on AI	https://www.oecd.org/going-digital/ai/principles/?utm_content=buffer018db&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer
OECD Council Recommendation on AI	https://legalinstruments.oecd.org/api/print?ids=648&lang=en
G20 Ministerial Statement on Trade and Digital Economy	https://www.mofa.go.jp/files/000486596.pdf
Fórum de Governança da Internet	
Declaração de Toronto: Protegendo os Direitos à Igualdade e à Não-Discriminação em Sistemas de Aprendizado por Máquinas (2018)	http://www.intgovforum.org/multilingual/sites/default/files/webform/toronto-declaration-final.pdf
Future of Life Institute	
Future of Life Institute - Asilomar AI Principles 2017	https://futureoflife.org/ai-principles/
FAT/ML	
FAT/ML - Principles for Accountable Algorithms and Social Impact Statement for Algorithms	https://www.fatml.org/resources/principles-for-accountable-algorithms
The Public Voice Coalition	
The Public Voice Coalition - Universal Guidelines for AI (2018)	https://thepublicvoice.org/ai-universal-guidelines/

International Conference of Data Protection and Privacy Commissioners	
ICDPPC - Declaration on Ethics and Data Protection in AI 2018	http://globalprivacyassembly.org/wp-content/uploads/2018/10/20180922_ICDPPC-40th_AI-Declaration_ADOPTED.pdf
Nova Zelândia	
Indigenous Protocol and AI Position Paper 2020	https://spectrum.library.concordia.ca/986506/
Algorithm Charter for Aotearoa New Zealand	https://data.govt.nz/manage-data/data-ethics/government-algorithm-transparency-and-accountability/algorithm-charter/
Austrália	
AI Ethics Principles	https://www.industry.gov.au/data-and-publications/building-australias-artificial-intelligence-capability/ai-ethics-framework/ai-ethics-principles
Uruguai	
Estrategia de Inteligencia Artificial para el Gobierno Digital	https://www.gub.uy/agencia-gobierno-electronico-sociedad-informacion-conocimiento/sites/agencia-gobierno-electronico-sociedad-informacion-conocimiento/files/documentos/publicaciones/Estrategia_IA%20-%20versi%C3%B3n%20espa%C3%B1ol.pdf
Argentina	
Plan Nacional de Inteligencia Artificial	https://ia-latam.com/wp-content/uploads/2020/09/Plan-Nacional-de-Inteligencia-Artificial.pdf
Colômbia	
Política Nacional para la Transformación Digital e Inteligencia Artificial	https://www.mintic.gov.co/portal/715/articles-107147_recurso_1.pdf
México	
Agenda Nacional Mexicana de Inteligencia Artificial	https://36dc704c-0d61-4da0-87fa-917581cbce16.filesusr.com/ugd/7be025_6f45f669e2fa4910b32671a001074987.pdf

INSTITUTO IGARAPÉ
a think and do tank

O Instituto Igarapé é um think and do tank independente focado nas áreas de segurança pública, climática e digital e suas consequências para a democracia. Seu objetivo é propor soluções e parcerias para desafios globais por meio de pesquisas, novas tecnologias, comunicação e influência em políticas públicas. O Instituto trabalha com governos, setor privado e sociedade civil para desenhar soluções baseadas em dados. Fomos premiados como a melhor ONG de Direitos Humanos no ano de 2018 e melhor think tank em política social pela Prospect Magazine em 2019.

Instituto Igarapé

Rio de Janeiro - RJ - Brasil
Tel/Fax: +55 (21) 3496-2114
contato@igarape.org.br
facebook.com/institutoigarape
twitter.com/igarape_org

www.igarape.org.br

Direção criativa e layout
Raphael Durão - STORMdesign.com.br

ISSN 2359-0998

www.igarape.org.br

INSTITUTO IGARAPÉ
a think and do tank