COLARES
ADVOGADOS

Inteligência Artificial:

Desenvolvimento, Proteção Jurídica e Regulação no Brasil

Outubro 2025

Inteligência Artificial:

Desenvolvimento, Proteção Jurídica e Regulação no Brasil

COLARES ADVOGADOS

COLARES
ADVOGADOS

1

IA e seus vários **desafios jurídicos**

1 IA e seus vários **desafios jurídicos**

COLARES ADVOGADOS

Há muito que se falar sobre Inteligência Artificial - não é à toa que o assunto parece estar por toda a parte. A IA tornou-se parte concreta do cotidiano de pessoas e empresa, desde as recomendações automáticas nos nossos e-mails às decisões mais complexas que moldam negócios, produtos e relações humanas. Esse avanço, embora fascinante, traz consigo vários debates sobre os impactos e repercussões da IA na vida humana, em suas diferentes esferas.

Neste e-book, reunimos reflexões e aprendizados construídos ao longo de nossa experiência assessorando empresas tradicionais e de base tecnológica que já convivem, na prática, com a IA — seja desenvolvendo modelos próprios, seja incorporando soluções de terceiros aos seus processos.

Apesar dos desdobramentos da IA serem muitos, aqui escolhemos focar em alguns dos principais debates e desafios relativos às repercussões jurídicas do uso, desenvolvimento ou exploração da IA por empresas, com o objetivo de oferecer uma visão clara e acessível sobre eles.

Falar em "Inteligência Artificial" é falar de múltiplas tecnologias, métodos e aplicações — cada uma com seu grau de complexidade e impacto.

Quanto maior o volume de dados processados, a autonomia da solução e o potencial impacto nos indivíduos, maiores também os riscos e as responsabilidades envolvidas.

O Brasil ainda caminha rumo a um marco regulatório próprio — o Projeto de Lei nº 2.338/2023, conhecido como Marco Legal da IA, inspirado no AI Act europeu. Mas mesmo antes da aprovação dessa norma, diversas legislações já recaem sobre as empresas, em especial a Lei Geral de Proteção de Dados Pessoais (LGPD) e o Código de Defesa do Consumidor. A própria Autoridade Nacional de Proteção de Dados (ANPD) vem se posicionando sobre tópicos ligados à Inteligência Artificial e esse foi um dos temas escolhidos para a sua agenda regulatória de 2025 – 2026.

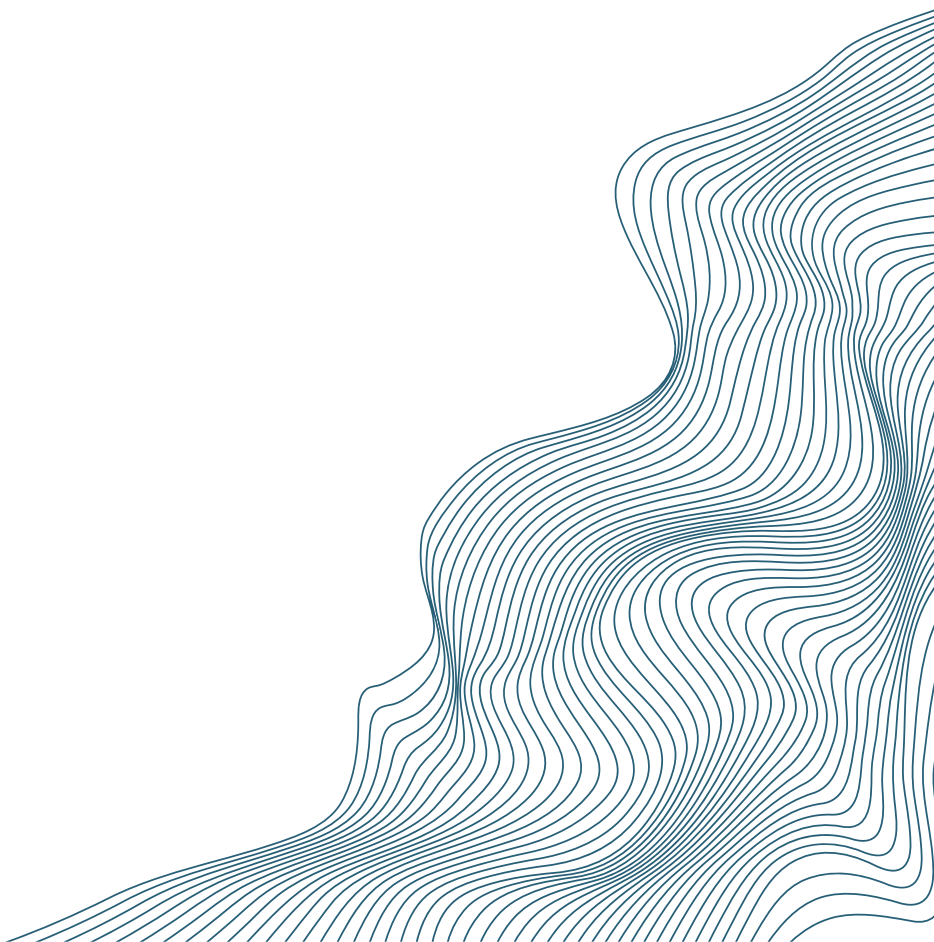
COLARES
ADVOGADOS

É preciso, portanto, se preparar para um ambiente regulatório mais rigoroso e estruturado. Ignorar essa realidade é expor-se a riscos jurídicos e reputacionais que podem comprometer a inovação e a sustentabilidade dos negócios.

Este material reflete a vivência prática do nosso time com organizações de diferentes perfis. Reunimos aqui alguns dos temas que, em 2025, mais têm desafiado empresas de diferentes setores — e que mais exigem respostas jurídicas sólidas, éticas e contemporâneas.

Esperamos que esta leitura ajude a compreender melhor os desafios apresentados e sirva como ponto de partida para novas discussões, inspirando o uso e a construção de soluções de Inteligência Artificial juridicamente seguras, sustentáveis e preparadas para o futuro.

Débora Vieira e **Rodrigo Colares**
Sócios do Colares Advogados



COLARES
ADVOGADOS

Modelos de IA como ativos estratégicos: formas de proteção

2 Modelos de IA como ativos estratégicos: formas de proteção

Quando falamos de modelos de IA, não estamos apenas diante de ferramentas que auxiliam o dia a dia das pessoas ou o aprimoramento de processos para diferentes organizações, mas também de ativos estratégicos que podem ser decisivos para a diferenciação competitiva de uma empresa, principalmente no mercado de tecnologia.

Tratar esses modelos como ativos significa reconhecer que eles podem ter alto valor econômico e estratégico e, por isso, precisam da proteção jurídica adequada para que a empresa assegure sua titularidade sobre eles e o direito de explorá-los, permitindo que seu potencial seja melhor aproveitado.

Ativos diferentes, proteções diferentes

Um modelo de IA é formado por diferentes elementos, e cada um deles está sujeito a um ou mais regimes jurídicos de proteção no Brasil. Uma boa estratégia de proteção de ativos de IA pode envolver e combinar diferentes regimes jurídicos para proteger diferentes aspectos da tecnologia.

(i) *Software:*

Software é protegido pela Lei de Direitos Autorais (Lei nº 9.610/98) e por lei própria, a Lei do Software (Lei nº 9.609/98). Seu código-fonte e demais elementos caracterizadores podem ser registrados junto ao Instituto Nacional da Propriedade Industrial (INPI), uma medida rápida e de baixo custo que funciona como importante prova de autoria — mas que não impede que terceiros resolvam o mesmo problema com um código diferente. Já se uma invenção é implementada por meio de *software*, é possível que ele possa ser patenteado, assegurando ao titular a exclusividade sobre as funções patenteadas. Sem qualquer prejuízo a essas duas formas de proteção, o *software* pode ainda ser protegido por disposições contratuais — talvez a forma de proteção mais comum —, como será detalhado mais à frente.

COLARES
ADVOGADOS

(ii) Algoritmos

Em um *software* podem existir diversos algoritmos, mas estes, por sua vez, não possuem tutela jurídica própria de acordo com a legislação brasileira. Ainda assim, enquanto parte de um *software*, algoritmos podem aproveitar as formas de proteção mencionadas acima.
Por exemplo: um ou mais algoritmos podem fazer parte de um software registrado junto ao INPI ou, se o *software* implementa uma invenção que pode ser patenteada. Além disso, algoritmos podem fazer parte de segredos de negócio ou informações confidenciais de empresas, os quais, como será visto adiante, podem ser protegidos pelas cláusulas contratuais adequadas.

(iii) Bases de dados

De forma geral, a Lei de Direitos Autorais brasileira estabelece que as bases de dados serão protegidas por direitos autorais quando houver originalidade na organização dos dados, ou seja, a forma de expressão da estrutura da base de dados pode estar resguardada, mas não o seu conteúdo. A proteção do conteúdo em si depende, mais uma vez, de disposições contratuais específicas que definam direitos e obrigações relacionados ao uso e à exploração da base de dados.

(iv) *Know-how* e informações confidenciais

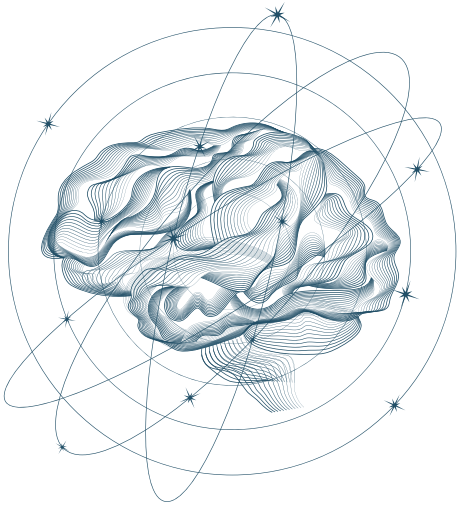
O *know-how* se refere a conhecimentos práticos e técnicos que oferecem uma vantagem competitiva à empresa, e não são necessariamente registrados ou patenteados. Já as informações confidenciais são aquelas que não são de conhecimento geral, dentro ou fora de uma organização, mas que também representam algum valor estratégico para a empresa. Tanto o *know-how* como as informações confidenciais podem ser protegidos por cláusulas contratuais de confidencialidade. Se protegidos contratualmente, bem como sujeitos a medidas de segurança adequadas (como, por exemplo, criptografia dos dados ou treinamentos internos), essas informações podem ser consideradas segredos de negócio, que recebem forma adicional de proteção legal, uma vez que a Lei da Propriedade Industrial (LPI) prevê que a sua violação pode configurar crime de concorrência desleal.

Ou seja, não há forma única de proteger um modelo de IA.

É preciso identificar quais os elementos relevantes envolvidos e, sempre que possível, combinar diferentes formas de proteção legal e contratual para mitigar ao máximo os riscos.

A propósito, ainda não existe, de acordo com a legislação brasileira, a possibilidade de uma IA ser considerada titular de direitos, autora de uma obra protegida por direitos autorais ou inventora de uma patente. O próprio INPI proferiu decisão, no processo de nº BR 11 2021 008931 4 A2, argumentando pela impossibilidade de indicação, pelo requerente de pedido de patente, de máquina dotada de IA como inventora da patente, havendo: "necessidade de edição de legislação específica, possivelmente antecedida pela celebração de tratados internacionais destinados a uniformizar o tratamento do tema".

Por isso, se os resultados gerados pela IA também são ativos relevantes para a empresa, é **fundamental identificar quais regimes de proteção jurídica são aplicáveis e considerá-los na definição da estratégia de propriedade intelectual da empresa.**



O **PL nº 2.338/2023,** que tramita no Congresso Nacional e propõe estabelecer o Marco Legal de IA no Brasil, em sua versão atual, não se debruça em questões relativas à propriedade intelectual de modelos de IA, nem sobre a titularidade de seus *outputs*.

Contudo, já existem ferramentas jurídicas, principalmente contratuais, que podem ser utilizadas para fomentar tanto a segurança quanto o desenvolvimento, licenciamento, venda ou outras formas de exploração de modelos de IA. **Os contratos são importantes ferramentas pelas quais as empresas têm, atualmente, a oportunidade de proteger juridicamente seus modelos de IA.**

COLARES
ADVOGADOS

Proteções contratuais no desenvolvimento

Em relação ao desenvolvimento de modelos de IA, seja ele realizado internamente, por colaboradores da empresa, ou externamente, por prestadores de serviço contratados, é imprescindível que os instrumentos contratuais adotados possuam cláusulas robustas de propriedade intelectual, que estabeleçam de quem será a titularidade de toda a PI relacionada ao modelo de IA que tenha sido ou esteja sendo desenvolvido, assim como suas respectivas versões, melhorias, *softwares*, algoritmos e quaisquer outras informações, inclusive confidenciais, ou elementos a ele relacionados.

Caso o modelo de IA seja desenvolvido por terceiros, é importante para a empresa contratante negociar claramente (e refletir os entendimentos em **cláusulas contratuais adequadas) de quem será a titularidade do modelo, dos *inputs* (dados de treinamento), dos *outputs* (respostas ou resultados produzidos pelo modelo) e até mesmo dos seus aprendizados futuros.**

Sempre que os ativos forem criados por terceiros, mas serão de titularidade da contratante, é essencial prever cláusulas de cessão adequadas. Já se os ativos desenvolvidos forem de titularidade da prestadora de serviços, é importante prever o seu licenciamento em favor da empresa contratante e dispor detalhadamente o que pode ou não ser feito com eles, se há exclusividade e quais as condições que regerão o licenciamento — com especial atenção às situações em que o licenciamento poderá ser encerrado.

Todas essas questões merecem atenção e discussão, desde as primeiras conversas até a finalização do contrato, para garantir que as expectativas das partes estejam alinhadas, uma vez que tais definições são estratégicas e podem impactar diretamente a vantagem competitiva da empresa contratante e sua capacidade de monetizar a tecnologia desenvolvida.

Também é essencial que os desenvolvedores do modelo, sejam internos ou externos, comprometam-se a não violar direitos de terceiros durante o desenvolvimento e apenas utilizem softwares de terceiros ou open-source com a ciência ou autorização da empresa contratante.

Nesse sentido, é importante atentar ao fato de que o uso de softwares *open-source* (ou abertos), comum em projetos de IA, ainda que possa reduzir custos e acelerar o desenvolvimento do modelo, envolve relevantes riscos que devem ser ponderados.

Merecem especial atenção, nesses casos, as licenças *copyleft*, que permitem o uso gratuito de *software open-source*, mas exigem que *softwares* delas derivados sejam distribuídos sob a mesma modalidade de licença (isto é, de forma também gratuita), o que pode obrigar a divulgação de códigos e comprometer a exclusividade de ativos estratégicos, além de inviabilizar modelos de negócio baseados em licenças de software pagas.

Além disso, o uso de softwares abertos também traz o risco de cometimento não intencional de infrações, já que parte do código disponibilizado pode conter trechos cuja origem não é clara ou que já possuem violações pré-existentes à sua utilização no modelo de IA, como a incorporação de modelos pagos ou licenciados indevidamente, por exemplo.

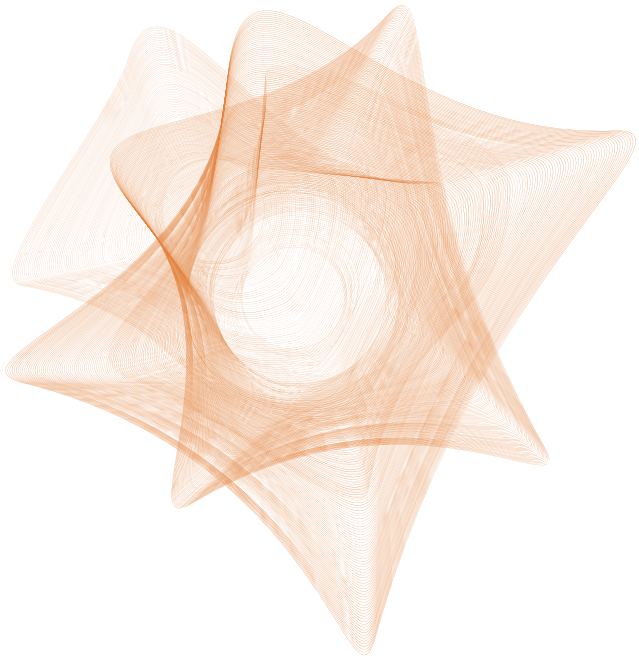
Por esses motivos, é importante entender o que é compatível ou necessário para o desenvolvimento do modelo de IA de acordo com o negócio, o modelo de exploração pretendido e o apetite de risco da empresa.

Caso componentes open-source sejam utilizados, é recomendável que empresas sempre analisem as licenças sob as quais esses componentes são disponibilizados e mantenham inventários atualizados de todos eles, buscando garantir o cumprimento rigoroso das condições de suas licenças. Esse tipo de documentação, inclusive, é frequentemente exigido em processos de *due diligence* ligados à captação de investimentos ou aquisição de empresas.

Proteções contratuais na exploração

Para além disso, para proteger os interesses da empresa que será a detentora dos direitos sobre o modelo de IA, é essencial estabelecer obrigações robustas de confidencialidade, por meio da assinatura de acordos de confidencialidade (também conhecidos como *Non-Disclosure Agreements* ou NDAs) ou inserção de cláusulas no contrato de trabalho ou de prestação de serviços celebrados junto aos desenvolvedores, bem como nas relações junto a parceiros ou clientes que venham a ter acesso ou utilizar a IA desenvolvida.

Enquanto for um ativo estratégico da empresa, o modelo de IA deve ser protegido em todo o seu ciclo de vida, inclusive durante o seu uso interno ou exploração junto a terceiros.



Nestes casos, é importante, ainda, que o contrato descreva e delimite os escopos e entregas, além de possuir disposições relacionadas à proteção de dados e segurança da informação. Caso a empresa seja a desenvolvedora de um modelo de IA que será fornecido a terceiros, cláusulas de não aliciamento podem ser importantes para proteger colaboradores estratégicos, que podem ser diretamente abordados pelos usuários da solução para o desenvolvimento de sistema igual ou similar.

Por outro lado, caso seja negociada entre as partes exclusividade em relação ao modelo de IA, disposições relativas à não concorrência podem fazer sentido para assegurar que o desenvolvedor não explore comercialmente o modelo nem desenvolva similar para um concorrente.

Em casos de contratos de licenciamento contínuo de um modelo de IA que exija suporte e manutenção, é importante para as duas partes estabelecer as métricas de acordos de nível de serviço (SLAs) a serem respeitadas pelo fornecedor do modelo de IA, melhor dimensionando as expectativas de ambas as partes e dando previsibilidade às consequências do descumprimento das métricas acordadas (multas ou descontos, por exemplo).

O outro lado: cuidados contratuais para quem contrata

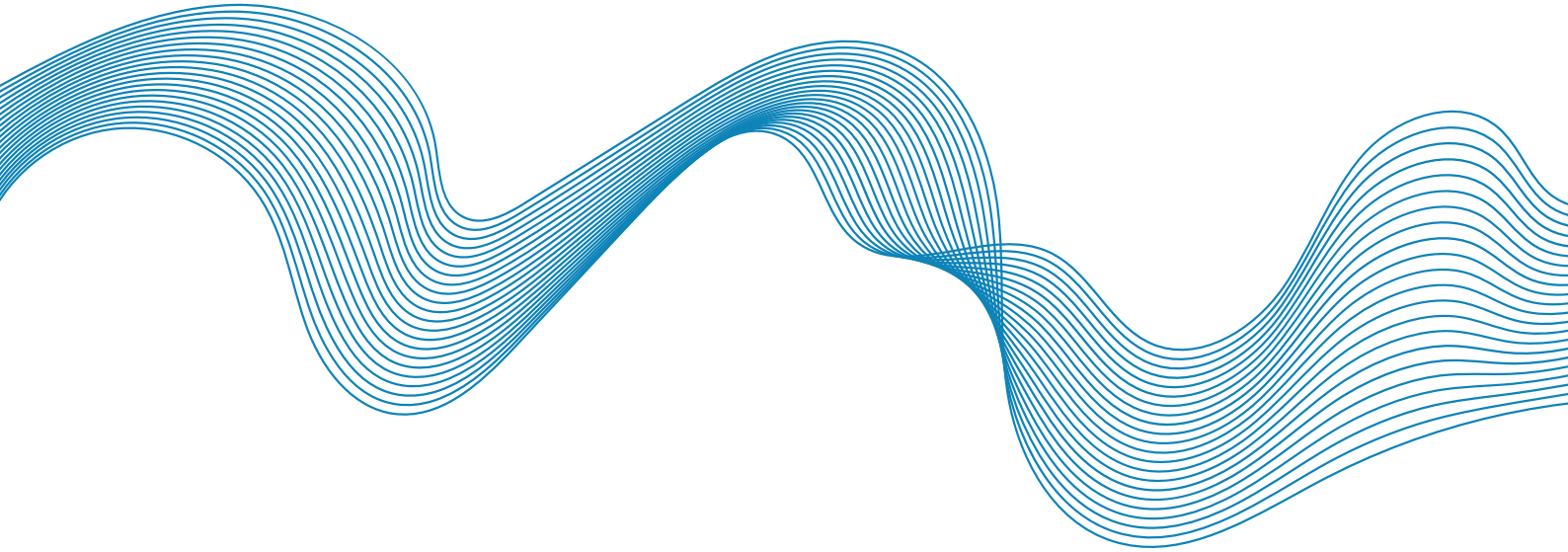
Outro ponto que merece atenção, especificamente para empresas que optam por contratar o licenciamento de ferramentas de IA, é a possibilidade de dependência tecnológica de soluções oferecidas sob regime de *SaaS* (*Software as a Service*), o *lock-in*.

Esse risco ocorre quando a empresa fica restritivamente vinculada a um único fornecedor, mesmo que este pratique valores muito altos ou não forneça os melhores serviços e/ou inovação, diante da existência de diversas barreiras técnicas, operacionais, contratuais ou financeiras que inviabilizam a migração para outra solução. Essa dependência pode comprometer a continuidade do negócio, limitar a capacidade de negociação, gerar custos mais elevados e dificultar o desenvolvimento da tecnologia da empresa.

Para mitigar esses impactos, é essencial negociar cláusulas flexíveis, buscando a sustentação da inovação do modelo de IA a longo prazo, como, por exemplo, cláusulas de portabilidade e interoperabilidade de dados que prevejam a possibilidade de migração assistida em caso de troca de fornecedor, assim como evitar disposições que prevejam penalidades severas por rescisão contratual antecipada.

Além disso, as empresas podem utilizar outros mecanismos tecnológicos para tentar se resguardar de situações de dependência, como a realização de backups periódicos ou utilização de APIs, ou utilizar múltiplos fornecedores (multi-vendor) para as diferentes funcionalidades essenciais ao negócio, de modo a evitar a dependência de um único fornecedor.

Em conclusão, modelos de IA são ativos valiosos, mas também carregam fragilidades jurídicas que devem ser um ponto de atenção, principalmente no que se relaciona à sua propriedade intelectual, demandando, entre outros pontos, uma estrutura contratual robusta, análises e diligências prévias e atenção aos objetivos e interesses do seu negócio. Um olhar completo para a PI do seu modelo de IA pode assegurar que modelos de IA sejam um diferencial competitivo, sustentável e seguro para a sua empresa.



COLARES
ADVOGADOS

3

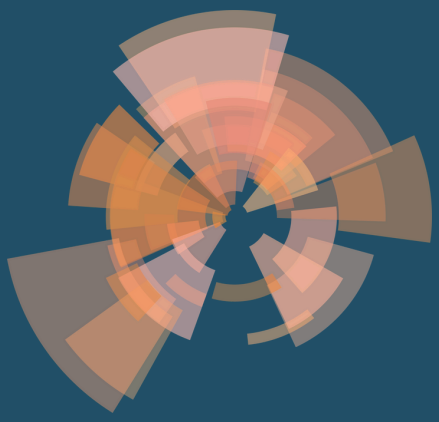
Desafios de proteção de dados no treinamento e uso de IA

COLARES
ADVOGADOS

3 Desafios de proteção de dados no treinamento e uso de IA

A expansão do uso de Inteligência Artificial (IA) traz consigo relevantes desafios para a proteção de dados pessoais. Se, por um lado, alguns modelos de IA demandam grandes volumes de informações para alcançar eficiência e precisão, por outro, a Lei Geral de Proteção de Dados (LGPD) impõe limites quanto ao tratamento desses dados.

Entre as questões de proteção de dados que mais têm preocupado as empresas que investem no desenvolvimento e uso de IA, estão a escolha de bases legais para justificar o tratamento de dados para fins de treinamento de sistemas de IA, os impactos do princípio da minimização diante dessa tecnologia que depende de grandes quantidades de dados, bem como as obrigações de transparência em relação às bases de dados usadas para treinamento de modelos de IA — que muitas vezes consistem em ativo valioso e diferencial competitivo.



Apesar desses e outros desafios de proteção de dados no contexto de IA parecerem, à primeira vista, sem solução, não é o caso. Enquanto o Marco Legal da IA não é aprovado no Brasil e regulamentado pelas futuras autoridades competentes, já há tendências internacionais que consolidam boas práticas sobre o tema, que podem — e devem — ser observadas desde já por empresas brasileiras.

Base legal para o tratamento de dados no treinamento de IA

Antes de se iniciar o tratamento de dados pessoais para o treinamento de qualquer modelo de IA, é preciso definir a base legal que o autoriza, uma vez que atividades de tratamento realizadas sem respaldo jurídico configuram tratamento ilícito de acordo com a LGPD e podem levar à responsabilização da empresa, inclusive com aplicação de multas expressivas.

COLARES
ADVOGADOS

Das 10 bases legais previstas na LGPD para o tratamento de dados pessoais, a maior parte se mostra inadequada para a finalidade de treinar modelos de IA, uma vez que não são compatíveis com esse fim. Aqui, é importante diferenciar a base legal utilizada para o treinamento daquela utilizada para o uso do modelo de IA: este último pode ter as mais diversas finalidades, permitindo mais flexibilidade na escolha de base legal adequada.

Dentre as bases legais disponíveis para o treinamento de modelos de IA, o **consentimento** — previsto nos artigos 7º, I, e 11º, I, da LGPD — é comumente cogitado, devido à percepção de controle que confere ao titular sobre os seus dados. Contudo, no contexto de tratamento de um volume significativo de dados pessoais para o treinamento de um modelo de IA, a adoção dessa base legal encontra barreiras significativas: a dificuldade de obtenção (e comprovação) do consentimento válido de um grande número de titulares e a possibilidade de revogação desse consentimento a qualquer tempo.

A obtenção de consentimento válido de um grande volume de titulares é extremamente complexa, pois é necessário garantir que cada titular receba informações claras, completas e específicas sobre o tratamento de seus dados — o que demanda algum nível de contato direto com cada titular, além de **mecanismos robustos de comunicação, transparência, registro e gestão de anuências, visto que um baixo grau de adesão pode inviabilizar a operação de treinamento do sistema.**

Além disso, a possibilidade de revogação do consentimento pelo titular a qualquer momento obriga o controlador a garantir a viabilidade de exclusão ou anonimização dos dados já coletados — medida que pode ser técnica ou financeiramente inviável e comprometer a continuidade do treinamento dos modelos de IA, afetando sua eficiência, precisão ou até exigindo a reexecução de etapas já concluídas.

Diante dessas limitações e da inaplicabilidade das demais bases legais previstas na legislação, o **legítimo interesse torna-se a base legal preferencial (ou a única viável) para o treinamento de modelos de IA, desde que não sejam utilizados dados sensíveis para esse fim**. Isso porque permite ao controlador tratar dados pessoais sem depender da anuência individual de cada titular, oferecendo maior estabilidade e segurança jurídica às operações.

Para que o emprego desta base legal seja válido, é necessário demonstrar que há um interesse — baseado em situações concretas — do controlador ou de um terceiro e que sejam observados rigorosamente os requisitos previstos pela LGPD. Para isso, é **altamente recomendável a realização de Avaliação de Legítimo Interesse (LIA)** para confirmar se existe equilíbrio entre os interesses do controlador e os direitos e liberdades fundamentais dos titulares, bem como que o tratamento respeita a legítima expectativa destes últimos.

Além disso, considerando que no tratamento de dados para modelos de IA frequentemente estão presentes os critérios de alto risco definidos pela ANPD, **também é necessária a elaboração do Relatório de Impacto à Proteção de Dados (RIPD)**, que identifica riscos envolvidos no tratamento de dados para determinada finalidade e define medidas técnicas e organizacionais que mitigam os riscos mapeados, podendo este relatório ser desenvolvido em conjunto com a Avaliação de Impacto Algorítmico (AIA), documentação que se tornará obrigatória para sistemas de IA de alto risco caso o Projeto de Lei nº 2.338/2023 (Marco Legal da IA) seja aprovado.

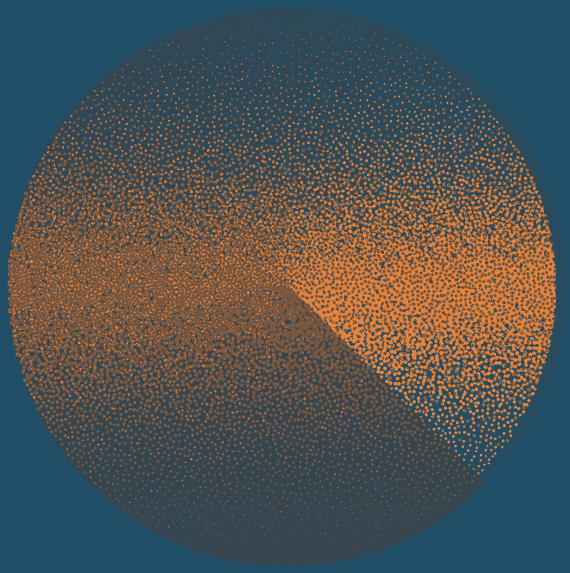
Princípio da minimização de dados e o treinamento de IA

Além do cumprimento dos requisitos legais para a escolha da base legal, é fundamental observar os princípios da LGPD, que devem orientar todo o tratamento de dados pessoais durante o treinamento de modelos de IA. Dentre eles, destaca-se o princípio da necessidade ou minimização.

Em particular, o princípio da necessidade ou minimização, previsto no artigo 6º, III, da LGPD, determina que apenas os dados estritamente necessários para atingir a **finalidade pretendida devem ser coletados e tratados, evitando excessos, tratamentos desnecessários e potenciais enviesamentos.**

COLARES
ADVOGADOS

No contexto do treinamento de modelos de IA, a aplicação do princípio da minimização torna-se especialmente desafiadora, sobretudo quando a operação depende do tratamento de grandes volumes de dados possivelmente provenientes de múltiplas fontes.



A discussão sobre o que é "estritamente necessário" no contexto de treinamento de IA não é fácil, considerando que não se está diante de apenas duas possibilidades claras (finalidade atingida vs. finalidade não atingida), mas sim de um **complexo gradiente**, no qual cada base de dados adicional pode significar uma maior qualidade do modelo.

No caso de IAs de propósito geral — isto é, aqueles modelos que não foram projetados para atingir somente uma finalidade, mas para desempenhar variadas funções com propósitos distintos —, a observância ao princípio da minimização é ainda mais difícil, pois a execução das suas múltiplas finalidades exige um treinamento realizado a partir de bases de dados mais extensas, e cada finalidade tem uma escala quase infinita de graus em que pode ser mais ou menos atingida.

Afinal, quanto maior e mais diversificado o conjunto de dados utilizado no treinamento do modelo, maior será a sua capacidade de aprender padrões complexos e gerar resultados mais precisos, o que gera uma aparente incompatibilidade com a restrição legal de limitar o tratamento ao mínimo necessário.

No plano internacional, diversas autoridades e organismos de proteção de dados têm emitido recomendações sobre o uso de dados pessoais no treinamento de modelos de IA que estão alinhadas aos princípios da LGPD, promovendo boas práticas relacionadas à minimização.

França | Commission Nationale de l'Informatique et des Libertés (CNIL)

A Commission Nationale de l'Informatique et des Libertés (CNIL), da França, afirma que, mesmo no caso de IAs de propósito geral, é possível delimitar, ainda que não absolutamente, as finalidades do sistema através da identificação prévia de suas capacidades que oferecem os maiores riscos, da descrição das funcionalidades excluídas na fase de desenvolvimento do sistema, ou ainda da indicação das condições de uso e dos casos de aplicação previstos. Dessa forma, é possível alcançar um nível de minimização satisfatório, assegurando que os dados tratados sejam proporcionais às finalidades esperadas.

Reino Unido | Information Commissioner's Office (ICO)

O Information Commissioner's Office (ICO), do Reino Unido, complementa esse entendimento ressaltando que o fato de alguns dados poderem vir a ser úteis no futuro não justifica, por si só, sua coleta, uso ou retenção. Ou seja, a expectativa de um eventual treinamento de modelos de IA não autoriza a manutenção dos dados coletados e tratados previamente. Quando o treinamento de um modelo de IA passar a constituir uma finalidade concreta, o controlador deverá realizar uma nova coleta, indicar a base legal adequada para o tratamento de dados, garantir o fácil exercício dos direitos dos titulares e informá-los claramente sobre a nova finalidade.

Adicionalmente, o ICO entende que para compatibilizar o princípio da minimização com os treinamentos de IA que utilizam dados em larga escala, podem ser adotadas soluções técnicas como:

> (i) a adição de perturbação ou ruídos às bases de dados, modificando discretamente os dados originais para diminuir a identificabilidade dos indivíduos, sem comprometer a utilidade geral dos dados para o treinamento;
>
> (ii) o uso de dados sintéticos, isto é, conjuntos de dados artificiais com padrões estatísticos semelhantes aos reais, reduzindo a necessidade de manipular dados pessoais; e
>
> (iii) o aprendizado federado, que treina modelos localmente em diferentes dispositivos ou bases de dados isoladas, sem a transferência de dados pessoais brutos entre eles.

COLARES
ADVOGADOS

Princípio da transparência e o conflito **com a proteção aos segredos comerciais**

Outro princípio da LGPD que merece especial atenção no contexto de treinamento e uso de IA é o da **transparência**. A LGPD o consagra como princípio fundamental em seu artigo 6º, VI, exigindo que as práticas de tratamento sejam comunicadas de forma clara e facilmente acessível aos titulares de dados, evitando informações obscuras ou em linguagem excessivamente técnica. A lei ainda reforça, em seu artigo 18, os direitos de acesso e à informação, permitindo que os titulares de dados solicitem informações como a confirmação da existência de tratamento, quais dados estão sendo tratados e para quais finalidades.

Apesar disso, as bases de dados utilizadas para treinamento de modelos de IA são frequentemente consideradas ativos estratégicos e divulgar informações precisas sobre quais bases são utilizadas poderia comprometer o valor econômico dos próprios modelos, além de revelar segredos comerciais das empresas.

Por isso, é fundamental encontrar um **equilíbrio entre a obrigação de transparência e a proteção do negócio**, adotando comunicações proporcionais, sem a necessidade de divulgar detalhes que comprometam a competitividade do negócio.

Autoridades e organismos de proteção de dados também têm orientado empresas sobre como alcançar esse equilíbrio. A CNIL, já mencionada, recomenda que **as organizações respeitem um intervalo temporal razoável** entre o momento em que os titulares são informados de que seus dados integram uma base de dados de treinamento e a efetiva utilização desses dados no treinamento de um modelo de IA. Isso garante que os titulares tenham tempo suficiente para exercerem seus direitos, inclusive o de oposição ao tratamento e solicitação da exclusão de seus dados.

Além disso, a CNIL orienta que essa informação seja acompanhada de detalhes claros sobre a origem dos dados — por exemplo:

> (i) se são provenientes de uma **base de dados pública**, é importante indicar o link para acesso; ou
>
> (ii) se são provenientes de um **corretor de dados (*data broker*)**, é importante indicar suas informações de contato —, condições de coleta e rotulagem, canais de comunicação com o controlador original e características do conjunto de dados (*dataset*) utilizado para o treinamento.

COLARES
ADVOGADOS

Ainda, o princípio da **explicabilidade**, já adotado na legislação de outros países e incorporado ao PL nº 2.338/2023, também é uma forma de complementar e dar efetividade prática ao princípio da transparência, garantindo que os titulares tenham condições de compreender de maneira adequada como seus dados influenciam as decisões dos sistemas e quais efeitos podem decorrer do seu uso, mesmo sem acesso integral às bases ou aos algoritmos por trás dos modelos.

Esperar a regulamentação **pode custar caro**

Apesar das obrigações de proteção de dados previstas na LGPD já estarem em vigor desde 2020, a sua observação em relação a modelos de IA — sobretudo quanto ao seu treinamento — ainda é relativamente recente.

A Autoridade Nacional de Proteção de Dados (ANPD) brasileira começa a se debruçar sobre o tema e deve emitir mais orientações a respeito no futuro próximo, até mesmo antes de um futuro Marco Legal da IA ser aprovado e entrar em vigor. Isso, contudo, não afasta as obrigações previstas na LGPD, que continuam sendo exigíveis e devem ser observadas por empresas também em relação aos modelos de IA.

Pela própria natureza dos modelos de IA e do seu treinamento, **decisões de negócio que contrariem princípios e obrigações de proteção de dados podem ser irreversíveis e custar bastante caro**, eventualmente comprometendo a própria viabilidade do modelo.

Aqui, é fundamental o diálogo entre times envolvidos no desenvolvimento e treinamento de modelos de IA e a equipe interna ou externa responsável por privacidade e proteção de dados, **atentando à construção de um modelo *privacy-by-design***.

COLARES
ADVOGADOS

Explicabilidade e revisão de decisões tomadas por IA

COLARES
ADVOGADOS

4 Explicabilidade e revisão de decisões tomadas por IA

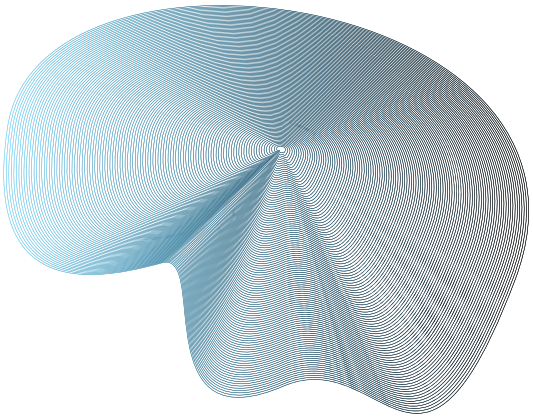
Com cada vez mais frequência, diferentes decisões de empresas, em seus mais variados departamentos, são delegadas a ferramentas de Inteligência Artificial — ou, pelo menos, fortemente influenciadas por elas.

Os possíveis impactos que essas decisões podem gerar, sobretudo em relação a pessoas, estão diretamente ligados à criação de direitos para que a pessoa afetada por uma decisão automatizada possa entendê-la, assim como os elementos nos quais ela se baseou, e, sempre que entender que a decisão foi injusta ou incorreta, solicitar a sua revisão.

Esses direitos — frequentemente chamados de **direito à explicação e à revisão de decisões automatizadas** — foram inicialmente previstos, no Brasil, pela Lei Geral de Proteção de Dados (LGPD), que seguiu uma tendência internacional de estabelecer tais direitos, mas não os detalhou o suficiente. Com a tramitação do Projeto de Lei nº 2.338/23, que deve criar o Marco Legal da IA, o assunto volta à tona e os desafios de explicar e revisar decisões de IA ganham forma mais definida.

Direito à **revisão**

A LGPD estabelece, em seu artigo 20, o direito do titular dos dados **a solicitar a revisão de decisões tomadas exclusivamente com base em tratamento automatizado de dados pessoais**, sempre que tais decisões afetarem seus interesses — o que inclui, por exemplo, decisões relativas à definição de perfil pessoal, profissional, de consumo ou de crédito. Vale destacar que, inicialmente, o texto da LGPD previa expressamente que essa revisão fosse realizada por **pessoa natural.**



Contudo, esse dispositivo foi posteriormente **suprimido**, restando apenas o direito de solicitar a revisão, sem que fique claro quem — ou o quê — deve realizá-la. A alteração legal foi alvo de críticas por parte de especialistas, que apontaram para o possível **esvaziamento deste direito caso a revisão de uma decisão tomada por IA seja realizada, novamente, por IA.**

Apesar da lacuna na LGPD, o entendimento predominante (e mais conservador) da doutrina e dos profissionais da área é de que a revisão deve, sim, ser **humana**. É nesse sentido, por exemplo, que o Regulamento Geral de Proteção de Dados (GDPR) da União Europeia — no qual a LGPD foi largamente baseada — prevê:

o titular de dados sujeito a uma decisão automatizada tem direito à intervenção humana.

O Projeto de Lei nº 2.338/2023, que atualmente tramita no Congresso Nacional, evidencia o direito do titular de dados a receber explicações claras sobre decisões, recomendações ou previsões feitas por sistemas de IA classificados como de alto risco, conforme critérios estabelecidos no próprio texto legislativo. Além disso, o projeto garante a possibilidade de o titular contestar essas decisões e pedir sua revisão, resolvendo a lacuna deixada pela LGPD e estabelecendo, explicitamente, que tal revisão deve ser conduzida por um humano.

No cenário europeu, o AI Act, regulamento sobre o uso seguro e responsável de sistemas de IA aprovado em 2024, já reforçava a necessidade de intervenção humana em decisões automatizadas de alto risco — caminho que o PL brasileiro acertadamente também deve seguir.

Uma outra importante questão pode ser aprendida a partir das discussões europeias, enquanto não há orientações nacionais sobre o assunto: para se considerar que houve revisão humana, essa intervenção no processo deve ser significativa, ou seja, o revisor não pode simplesmente confirmar o resultado algorítmico. Em vez disso, deve dispor de informações e autonomia suficientes para reavaliar o caso e, conforme seu próprio entendimento, alterar o teor da decisão.

Isto é, as empresas que desenvolvem e utilizam IA devem estar preparadas para oferecer não apenas uma revisão formal e mecânica, na qual o humano apenas legitima o resultado do algoritmo, sem examiná-lo criticamente ou contar com os meios para modificá-lo, mas sim uma revisão que leve à possibilidade concreta de alteração da decisão.

Explicabilidade

A LGPD também impõe ao controlador dos dados, aquele a quem competem as decisões referentes ao tratamento, a obrigação de fornecer **informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados** para a tomada da decisão automatizada, ressalvados os segredos comercial e industrial. Ou seja: a pessoa diretamente afetada por uma decisão automatizada tem direito a entender, de forma transparente, como ela foi tomada.

É da ANPD, inclusive, a competência para realizar auditoria em caso de recusa injustificada do controlador em fornecer essas informações, especialmente quando o controlador invoca, de forma abusiva, segredo comercial ou industrial.

O direito de receber tais explicações — comumente referido como o direito à explicação (ou explicabilidade) — foi criado pela LGPD, mas não recebeu detalhamento sobre o grau de explicação que deve ser dado, seu nível de profundidade ou tecnicidade.

Entretanto, tendo em mente o princípio da transparência que rege a LGPD, o entendimento predominante entre os profissionais da área é de que o tratamento de dados pelo modelo de IA deve ser explicado de modo que assegure a compreensão, pelo titular dos dados, dos fundamentos que levaram à decisão, inclusive para que ele possa eventualmente contestá-la.

Ou seja: o direito à explicação e o direito à revisão estão fortemente ligados e a forma como o primeiro é garantido pode impactar diretamente a efetividade do segundo.

A respeito, o Projeto de Lei nº 2.338/2023, à semelhança do AI Act europeu, oferece um pouco mais de detalhamento ao criar expressamente o direito à explicação sobre a decisão, recomendação ou previsão feita por IA, estabelecendo que as explicações incluam informações suficientes, adequadas e inteligíveis, e que sejam apresentadas em linguagem simples, acessível e que facilite à pessoa compreender o resultado da decisão ou previsão. Apesar de ainda haver lacunas que serão objeto de regulamento pelas autoridades competentes, o PL já torna claro que a explicação deve ser adequada ao entendimento do titular — e não excessivamente técnica, de forma que apenas especialistas conseguiriam compreender os fundamentos da decisão ou previsão.

Além disso, o PL prevê que um grau de transparência, explicabilidade e auditabilidade do sistema de IA que dificulte significativamente o seu controle ou supervisão pode ser um critério para que sistemas sejam considerados de alto risco, o que atrai automaticamente para seus desenvolvedores e utilizadores obrigações legais muito mais severas.

Panorama atual brasileiro

As lacunas deixadas pela LGPD sobre os direitos à revisão humana e à explicabilidade, relativos a decisões automatizadas, não afasta a necessidade de adoção de mecanismos, pelas empresas, para garantir o exercício desses direitos — pelo contrário, reforça a importância de atentar aos avanços regulatórios, legislativos e também a posicionamentos dos tribunais para melhor entender como suas obrigações se darão na prática.

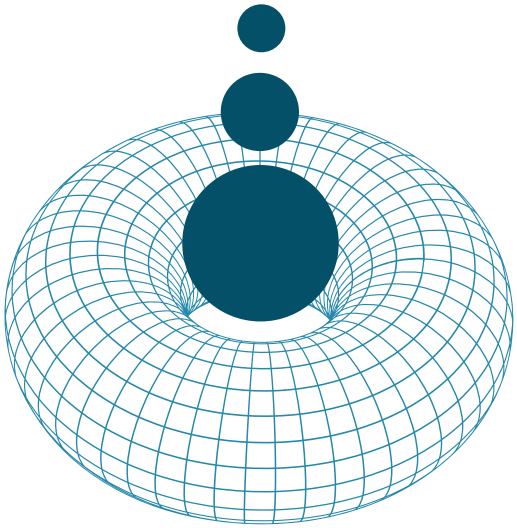
O Superior Tribunal de Justiça (STJ), ao julgar o Recurso Especial nº 2.135.783/DF, analisou um caso paradigmático envolvendo o descredenciamento de um motorista de aplicativo com base em uma decisão tomada de forma exclusivamente automatizada, sem prévia notificação do motorista ou possibilidade de revisão. **O STJ entendeu que tais decisões, por envolverem dados pessoais, estavam sujeitas à LGPD, conferindo ao titular o direito de ser informado acerca dos fundamentos do descredenciamento e de requerer sua revisão, em conformidade com o princípio da transparência e com o disposto no artigo 20 da lei.**

Embora o STJ não tenha declarado expressamente a obrigatoriedade de revisão humana, o julgamento evidencia que as exigências de explicabilidade e revisão já são aplicáveis na prática. Esse precedente também destaca a conexão entre os dois direitos: **a disponibilização de uma explicação clara e fundamentada sobre os critérios da decisão é fundamental para o exercício pleno do direito de revisão.**

Em paralelo, a ANPD já promoveu tomada de subsídios sobre decisões automatizadas, incluindo questões relativas à explicabilidade e à revisão, em preparação à elaboração de regulamento sobre o tema.

Os direitos dos titulares (incluindo aqueles relativos a decisões automatizadas) e Inteligência Artificial estão entre os temas prioritários elencados na agenda regulatória da ANPD para o biênio 2025 - 2026. **Isso significa que a regulamentação do assunto pela ANPD pode ocorrer antes mesmo que o Projeto do Marco Legal da IA seja aprovado no Congresso Nacional.**

Considerando o texto atual do Marco Legal da IA, ainda em tramitação, mas bastante alinhado às tendências internacionais, parece prudente que as empresas se preparem, desde já, para viabilizar a revisão humana das decisões automatizadas de maior impacto e fornecer explicações claras e acessíveis sobre os critérios e fundamentos que levam às decisões tomadas por seus sistemas.



Isso requer a criação de processos internos sólidos de governança algorítmica, a capacitação de equipes dedicadas à revisão e a implementação de fluxos de atendimento que efetivamente assegurem o exercício dos direitos à explicação e revisão pelos titulares de dados.

COLARES
ADVOGADOS

5

IA e Viés Algorítmico: entre a inovação e a possível discriminação

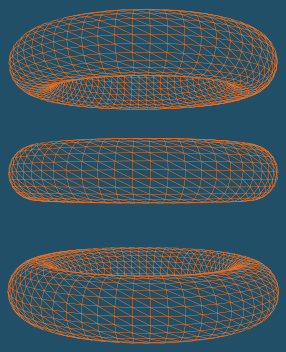
COLARES
ADVOGADOS

5 IA e Viés Algorítmico: entre a inovação e a possível discriminação

A busca por eficiência e objetividade tem levado empresas de todos os setores a adotarem a Inteligência Artificial (IA) na tomada de decisões. A ideia de que um algoritmo pode não apenas ser mais rápido, mas também superar as inconsistências do julgamento humano pode ser atrativa, mas a idealização de uma tecnologia supostamente “neutra” esconde uma realidade complexa.

Longe de serem imparciais, sistemas de IA são criados por seres humanos, que, conscientemente ou não, imputam no seu desenvolvimento suas próprias crenças, opiniões e valores.

Como resultado, muitos sistemas de IA acabam por refletir e amplificar os preconceitos estruturais da sociedade, criando um verdadeiro **"paradoxo da objetividade"**, no qual a IA sistemicamente reproduz injustiças com uma legitimidade técnica que torna ainda mais difícil de contestá-las.



Diante desses riscos, o cenário regulatório brasileiro tende a determinar a adoção de medidas de mitigação de vieses. O **Projeto de Lei nº 2.338/2023**, atualmente em tramitação no Congresso Nacional, **está estabelecendo um caminho que exigirá mais transparência e supervisão humana, principalmente em sistemas de "alto risco"**, uma categoria criada pelo PL para identificar aqueles sistemas de IA usados para determinadas finalidades e contextos (indicadas na lei ou pelas autoridades competentes) que, pela probabilidade e gravidade dos impactos adversos que podem causar às pessoas afetadas por eles, atraem obrigações muito mais severas para as empresas que os desenvolvem ou os utilizam.

Em paralelo, a Autoridade Nacional de Proteção de Dados (ANPD) também tem se debruçado sobre o tema de IA e decisões automatizadas e deve regulamentá-lo, inclusive no que diz respeito aos vieses algorítmicos, no futuro próximo.

COLARES ADVOGADOS

De onde vem o **viés algorítmico?**

Quando falamos em "viés algorítmico", não estamos nos referindo a uma falha técnica isolada ou a um bug aleatório. Trata-se de um problema sistemático e persistente, que leva um sistema de IA a produzir resultados injustos, privilegiando um grupo de pessoas em detrimento de outro. Essa distorção não surge do nada. Ela é o resultado de falhas que ocorrem em todo o processo de criação e treinamento da tecnologia, fazendo com que ela não apenas reproduza, mas muitas vezes amplifique as desigualdades e os preconceitos do mundo real. Para entender como isso acontece, é preciso olhar para as suas múltiplas origens.

Vieses originados nos dados

Viés de Seleção

A fonte mais evidente de viés reside nos próprios dados utilizados para treinar os modelos de aprendizado de máquina. O chamado **viés de seleção ou amostragem** ocorre quando o conjunto de dados não representa adequadamente a diversidade da população sobre a qual o sistema irá operar. Um exemplo notório é o de sistemas de reconhecimento facial que, treinados predominantemente com imagens de indivíduos brancos, apresentam taxas de erro significativamente maiores ao analisar rostos de pessoas negras, com "falsos-positivos" — o que já teve consequências drásticas, resultando, inclusive, na **prisão de inocentes que foram equivocadamente "reconhecidos" como os suspeitos de atividades criminosas**.

Viés Histórico

Já o **viés histórico ou de preconceito** surge quando os dados usados para o treinamento do modelo, embora representativos, refletem desigualdades sociais passadas. Foi o que ocorreu, por exemplo, com a ferramenta de recrutamento da Amazon, que, ao ser treinada com currículos recebidos ao longo de uma década, "aprendeu" que o perfil de sucesso na indústria de tecnologia era majoritariamente masculino, passando a penalizar candidaturas que continham termos associados ao universo feminino.

COLARES ADVOGADOS

Vieses no design do sistema

O processo de desenvolvimento também é uma fonte crítica de vieses.

Viés Cognitivo

O **viés cognitivo** dos próprios desenvolvedores pode influenciar, de forma inconsciente, a arquitetura do modelo, a seleção de variáveis e a definição de métricas de sucesso.

Viés de Rotulagem

Intimamente ligado a isso está o **viés de rotulagem**, que surge durante a preparação dos dados, quando humanos classificam e rotulam os dados que serão usados no treinamento do modelo de forma inconsistente ou subjetiva, incorporando seus próprios preconceitos e opiniões no processo e levando o sistema a reconhecê-los da mesma forma.

Viés de Ponduração

Por fim, o **viés de ponderação** ocorre quando os desenvolvedores atribuem pesos inadequados a determinadas variáveis, conforme aquilo que pessoalmente consideram ser mais ou menos relevante, conferindo-lhes uma importância desproporcional na tomada de decisão do algoritmo, o que pode levar a resultados distorcidos.

COLARES
ADVOGADOS

Vieses de proxy e de avaliação

Alguns vieses podem ser mais difíceis de identificar à primeira vista. O **viés de proxy**, por exemplo, surge quando o sistema decide com base em uma variável aparentemente neutra, mas que, na prática, funciona como um substituto (proxy) para uma característica sensível protegida por lei. É o que ocorre, por exemplo, com um algoritmo de concessão de crédito que utiliza o CEP como um fator de risco, o que pode, indiretamente, discriminar indivíduos com base na sua raça e na sua classe socioeconômica, uma vez que pode existir uma forte correlação entre local de moradia e esses atributos.

Viés de Proxy

Viés de Avaliação

Por fim, outro viés relevante e que frequentemente é utilizado para "legitimar" os outros tipos de vieses mencionados acima é o **viés de avaliação**, que se manifesta quando as métricas para validar o desempenho do modelo são inadequadas.

Por exemplo, se o desempenho da IA é medido apenas com base na sua acurácia geral (quantidade de decisões corretas dentre o total de decisões tomadas), podem ser ignorados indícios de vieses que seriam percebidos a partir de recortes específicos. Um sistema pode apresentar alta acurácia geral, mas, ao mesmo tempo, ter um desempenho drasticamente inferior e prejudicial para um subgrupo minoritário específico, uma falha que métricas de avaliação simplistas não conseguem capturar, reforçando, consequentemente, a falsa crença de que o sistema é neutro, quando na verdade um viés impediu a identificação de outro.

COLARES
ADVOGADOS

Panorama jurídico

Ainda que o Brasil não possua, até o momento, um marco legal específico para a Inteligência Artificial, já existem mecanismos para coibir a discriminação algorítmica no país. A principal salvaguarda encontra-se, atualmente, na Lei Geral de Proteção de Dados Pessoais (LGPD), que, apesar de não tratar especificamente de IA, afeta diretamente os modelos que envolvem dados pessoais no seu treinamento ou uso.

O art. 6º, IX, da LGPD estabelece como um de seus princípios a não discriminação, proibindo categoricamente o tratamento de dados pessoais para fins "discriminatórios ilícitos ou abusivos".

Consequentemente, qualquer sistema de IA que, ao processar informações de indivíduos, produza resultados que prejudiquem injusta e sistematicamente grupos específicos com base em raça, gênero, origem ou outras características está em direta violação a um dos pilares da lei. Apesar de a LGPD não detalhar medidas específicas voltadas à mitigação de vieses algorítmicos, é um ônus do controlador (aquele a quem cabem as decisões do tratamento de dados) comprovar que está atuando em conformidade com a lei e todos os seus princípios.

Além disso, o futuro da regulação de IA no Brasil está sendo delineado pelo Projeto de Lei nº 2.338/2023, que se inspira em modelos internacionais como o AI Act europeu para estabelecer o Marco Legal de IA brasileiro. O PL aprofunda questões introduzidas pela LGPD, direito à explicação e à revisão humana de decisões, e a exigência de supervisão humana efetiva, garantindo que decisões de alto impacto não sejam deixadas exclusivamente a cargo das máquinas.

Além de reforçar o princípio da não discriminação já trazido pela LGPD, o PL estabelece também o direito à correção de vieses discriminatórios ilegais ou abusivos, sejam eles diretos ou indiretos — endereçando expressamente o problema sistêmico do qual estamos falando. De forma complementar, o texto legal estabelece também a obrigação de desenvolvedores e aplicadores (isto é, quem utiliza o sistema) de IA de adotarem medidas de governança para mitigar e prevenir vieses discriminatórios. Outra importante questão a ser observada no texto atual do PL é que o alto potencial de viés discriminatório deve ser um dos critérios para que autoridades competentes classifiquem sistemas de IA como de “alto risco”, o que significará uma carga de obrigações regulatórias mais severa.

COLARES
ADVOGADOS

Essa tendência legislativa reforça que medidas de governança, incluindo voltadas à mitigação de vieses, **devem ser implementadas por empresas que desenvolvem e utilizam IA desde já** — em atenção ao que já prevê a LGPD e em antecipação às exigências reforçadas que deverão vir com o novo Marco Legal.

Dicas de Governança:
Sugestões para mitigar os danos dos vieses

A mitigação de vieses algorítmicos não é um ato de correção pontual, mas um processo contínuo de governança que deve ser integrado a todo o ciclo de vida de um sistema de IA. Adotar a abordagem de *privacy by design* (privacidade desde a concepção) e, por extensão, *ethics by design* (ética desde a concepção) é fundamental para construir sistemas mais justos e reduzir riscos jurídicos e reputacionais.

Concepção e Design do sistema

A prevenção de vieses começa na fase de planejamento. É ideal montar equipes de desenvolvimento multidisciplinares e diversas, capazes de identificar pontos cegos que poderiam passar despercebidos quando todos têm a mesma perspectiva. Além disso, a definição cuidadosa do problema e das variáveis a serem utilizadas é essencial para evitar o uso de proxies discriminatórios — dados que parecem ser neutros, mas que possuem alta correlação com características sensíveis.

Curadoria e preparação dos dados

A qualidade dos dados de treinamento é determinante. A implementação de processos claros e auditáveis para a rotulagem dos dados é vital para minimizar a subjetividade e os preconceitos inconscientes dos anotadores humanos, assim como a utilização de bases de dados diversas, a fim de evitar a sub-representação de determinados grupos.

COLARES
ADVOGADOS

Validação e testes do modelo

A avaliação de um modelo não deve se limitar a métricas de assertividade geral (como a acurácia), mas também buscar observar se os níveis de assertividade são consideravelmente diferentes entre grupos distintos (por exemplo, entre homens e mulheres ou pessoas brancas e negras), o que pode sugerir que há um forte viés discriminatório no modelo.

O uso de **ferramentas de explicabilidade**, que ajudam a entender a lógica por trás das decisões do modelo, é um passo importante para identificar e corrigir correlações indesejadas.

Monitoramento contínuo

Após a implementação de novas medidas e processos, a governança continua. Para decisões de alto impacto, é indispensável manter um sistema de supervisão humana (human-in-the-loop), que permita a revisão e a reversão de decisões automatizadas. O monitoramento contínuo do desempenho do modelo em um ambiente real, em que pessoas reais são afetadas, é crucial para detectar desvios e vieses que possam surgir com o tempo. Finalmente, devem ser estabelecidos canais de feedback claros e acessíveis para que os indivíduos afetados possam contestar resultados e solicitar reavaliações, em conformidade com o direito de revisão previsto na LGPD.

O viés algorítmico é um risco intrínseco a sistemas baseados em *machine learning*, que pode ter diferentes causas e que é capaz de transformar ferramentas de inovação em mecanismos de discriminação. Como suas origens são diversas e o problema é sistêmico, as medidas de mitigação também devem ser sistêmicas e acompanhar todo o ciclo de vida do modelo de IA. O estabelecimento de uma governança de IA, com ou sem o apoio de especialistas externos, além de mitigar o risco específico ligado a vieses discriminatórios, viabiliza o cumprimento de forma mais assertiva de toda a rede de obrigações já existentes sob a LGPD (não se limitando a elas), e que deve ser intensificada com a iminente regulação específica da IA no Brasil.

Além disso, a implementação de uma governança de IA robusta ultrapassa a mera conformidade legal e se torna uma vantagem estratégica. Empresas que adotam proativamente medidas para mitigar vieses, auditando seus dados, diversificando suas equipes e garantindo a supervisão humana, não apenas constroem produtos mais justos e precisos, mas também fortalecem a confiança de seus clientes e se posicionam de forma resiliente diante de um cenário regulatório em evolução.

COLARES
ADVOGADOS

6

Cuidados na contratação e uso de ferramentas de IA

6 Cuidados na contratação e uso de ferramentas de IA

A adoção e o uso de ferramentas baseadas em Inteligência Artificial têm ocupado um lugar cada vez mais frequente nas atividades de empresas e seus colaboradores.

Com a crescente oferta de soluções "prontas para uso", incluindo aquelas adaptadas às necessidades de diferentes setores econômicos e departamentos, a delegação de tarefas à IA tornou-se uma realidade comum.

A facilidade de acesso a soluções prontas – como APIs, plataformas SaaS (*software as a service*) e funcionalidades de automação já integradas a sistemas utilizados pelas empresas – leva muitas equipes e colaboradores a testarem e adotarem essas ferramentas, efetivamente inserindo a IA nos processos da empresa, frequentemente sem que os tomadores de decisão estejam cientes de que sistemas de IA estão sendo utilizados – e, muito menos, dos riscos envolvidos.

O contexto de aplicação de cada solução sem dúvida impactará o grau do risco envolvido, mas temos observado alguns principais pontos de atenção comuns a diferentes contextos e áreas de negócio:

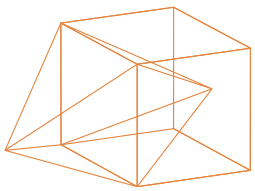
Desconhecimento dos termos da licença

Os termos sob os quais o sistema de IA é licenciado – de forma gratuita ou paga – descrevem questões cruciais relativas ao seu uso. Por exemplo, aqui podem estar contidas restrições para o uso comercial ou para certas aplicações. Além disso, devem estar contidas nesse documento as responsabilidades assumidas pelo fornecedor do sistema diante das obrigações já existentes relativas ao uso de IA, a exemplo da garantia da explicabilidade e mitigação de viés (tópicos abordados nos capítulos anteriores).

COLARES
ADVOGADOS

Uso dos dados para treinamento

Versões gratuitas de ferramentas generativas de IA (mas não apenas elas) normalmente utilizam os dados inseridos pelo usuário para treinamento do próprio modelo.
Caso informações confidenciais ou dados pessoais sejam fornecidos à solução, é possível que ela eventualmente utilize esses dados em respostas a outros usuários – uma violação a regras de confidencialidade que o colaborador pode nem saber que está cometendo.



Alucinações, vieses e respostas pouco confiáveis

Sistemas de IA não são neutros. Suas respostas podem conter vieses, dados desatualizados, erros crassos e até mesmo informações completamente inventadas. A utilização responsável de IA no trabalho – incluindo a formulação de prompts adequados e a verificação de respostas – deve ser ensinada e estimulada na empresa.

Delegação de decisões

A utilização de ferramentas de IA para automatizar tarefas corporativas pode levar a riscos particularmente elevados quando decisões significativas são delegadas à IA (ou fortemente influenciadas por ela). Por exemplo, decisões quanto à contratação de pessoas em processos seletivos ou avaliações de performance precisam ter critérios claros e a empresa deve ser capaz de fundamentá-las e explicá-las de modo transparente caso questionada – o que nem sempre é possível se a decisão tiver sido influenciada, ainda que parcialmente, por uma solução de IA.

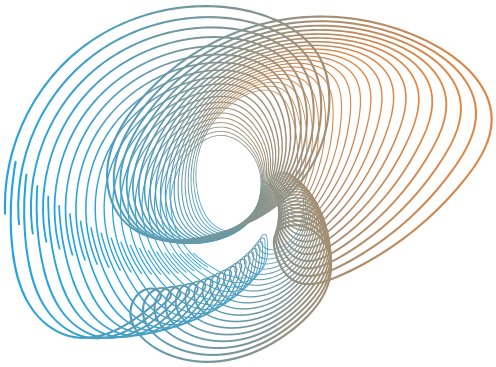
COLARES
ADVOGADOS

Mitigando riscos: mecanismos de prevenção

A mitigação destes e de outros riscos passa, em nosso entendimento, por uma combinação de medidas de conscientização e de governança interna, começando pelo reconhecimento de que, sim, os colaboradores de uma empresa provavelmente estão usando recursos baseados em IA. E, na ausência de orientações a respeito, este uso pode acontecer de várias formas – a maioria delas inadequada. Aqui vão algumas sugestões para lidar com o assunto:

Conscientização geral

Promover um **treinamento** básico, voltado a **todos os colaboradores**, sobre os riscos e as boas práticas no uso de IA na empresa. Essa é uma excelente oportunidade para promover uma conversa transparente sobre o assunto e mapear quais departamentos já usam recursos de IA – e para quais finalidades.



Treinamentos específicos

A depender dos usos feitos por diferentes departamentos e do nível de risco envolvido em cada caso, **é recomendável promover treinamentos específicos para times que façam maior uso de IA e/ou cujos usos geram maior risco para a empresa**.

Nessas formações, seriam discutidos **casos práticos**, como e em que situações soluções de IA poderiam ser utilizadas, considerando as necessidades específicas de cada área e os riscos envolvidos.

COLARES
ADVOGADOS

Contratação de ferramenta corporativa

Se soluções de IA são amplamente utilizadas por colaboradores da empresa, é recomendável avaliar a possibilidade de contratação de **licença corporativa** de uma ou de algumas soluções específicas como alternativa à utilização de versões gratuitas de várias ferramentas diferentes, sem que a gestão tenha visibilidade ou controle de seu uso.

Licenças corporativas normalmente oferecem melhores condições de confidencialidade e controle, como a não utilização dos dados imputados para treinamentos do modelo.

Em todo caso, a **análise dos termos da licença é fundamental, o que reforça a importância do próximo item.**

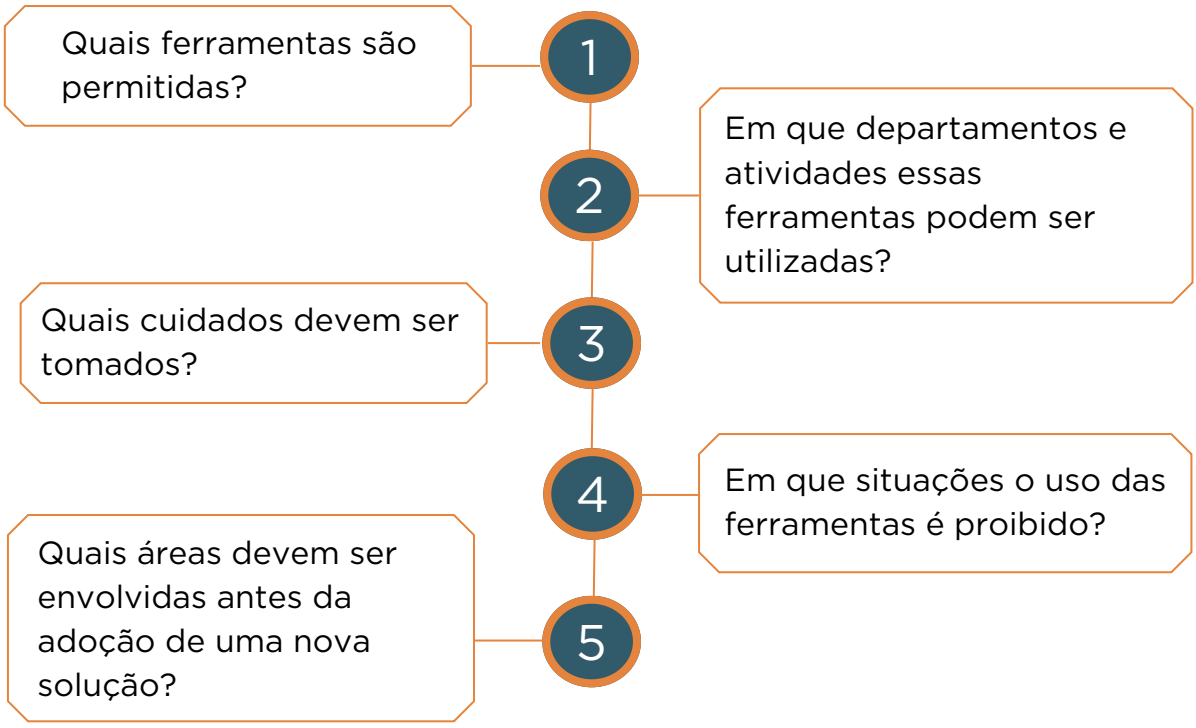
Envolvimento do time jurídico

A contratação (gratuita ou paga) de ferramentas de IA para uso no ambiente corporativo deve passar por uma avaliação jurídica, incluindo a análise dos termos da licença, das questões pelas quais o fornecedor da solução se responsabiliza (ou não) e de eventuais limitações relevantes para o uso que se pretende fazer dela.

A análise dos termos da licença é essencial para que a alta gestão decida quais riscos assumir e quais regras estabelecer.

Política interna de uso de IA

À medida que o assunto ganha maturidade na empresa, é recomendável estabelecer uma política interna que formalize as orientações e diretrizes corporativas para o uso da IA.



Essas e outras questões podem ser esclarecidas pela política interna, que também funciona como um instrumento de conscientização para novos colaboradores.

Considerando todos os ganhos proporcionados pelo uso de IA e o desejo de aumentar produtividade e eficiência, não nos parece aconselhável, ou mesmo viável, tentar frear a adoção de sistemas de Inteligência Artificial dentro de empresas. Assim como já se discutiu em relação a outras tecnologias e até mesmo ao uso de dados pessoais, entendemos que o caminho mais adequado para abordar o uso de IA nas empresas passa pela conscientização das equipes, identificação e mitigação de riscos e tomada de decisões informadas e alinhadas ao negócio.

Diante de um cenário regulatório ainda em construção, mas com uma clara tendência à criação de obrigações legais e responsabilidades para quem desenvolve ou utiliza ferramentas de IA, fomentar uma cultura de utilização responsável dessa tecnologia, desde já, é uma estratégia para minimizar riscos (atuais e futuros) e preparar a empresa para um processo de adequação à regulação nacional de IA. Empresas que utilizam IA com responsabilidade hoje estarão mais preparadas — e mais competitivas — no cenário regulatório de amanhã.

COLARES
ADVOGADOS

7

Marco Legal de IA no Brasil: o que esperar – e o que fazer desde já

7 Marco Legal de IA no Brasil: o que esperar – e o que fazer desde já

COLARES ADVOGADOS

À medida que a Inteligência Artificial se torna mais presente na realidade das pessoas e empresas, e gera impactos reais de diferentes naturezas, inclusive jurídicos, o Brasil caminha em direção ao estabelecimento de uma regulação específica do tema.

Atualmente, tramita no Congresso Nacional brasileiro o Projeto de Lei nº 2.338/2023, também já chamado de Marco Legal da Inteligência Artificial, o qual se destacou entre outros projetos apresentados ao propor uma regulação abrangente sobre IA, concentrando hoje os principais debates legislativos sobre o assunto. O PL já foi objeto de audiências públicas, importantes instrumentos de participação da sociedade e de especialistas de diferentes setores, e conta atualmente com estrutura robusta, que toca vários aspectos relevantes do desenvolvimento e uso de IA.

O Marco Legal da IA brasileiro, em seu texto atual, guarda similaridades com referências internacionais, em especial com o AI Act adotado em 2024 pela União Europeia, mas também acrescenta alguns novos elementos, decorrentes de aprendizados a partir da experiência europeia e das discussões promovidas no processo legislativo nacional. O projeto ainda está em tramitação e o seu texto ainda pode sofrer novos ajustes, mas já há fortes indícios quanto à direção que a regulação brasileira tende a seguir.

Uma das principais características do projeto brasileiro, à semelhança do modelo europeu, é o **modelo baseado na classificação de riscos**.

Como já visto, em razão da transversalidade da IA, os diferentes contextos e finalidades de seu uso podem levar a níveis de risco significativamente diferentes, de modo que não faria sentido aplicar o mesmo tratamento a todos os casos. Por isso, o PL estabelece níveis de obrigações proporcionais ao risco identificado em cada contexto, com obrigações mais severas quando o risco da solução de IA for mais alto, além de proibir usos que geram um nível de risco tão alto que é considerado inaceitável.

Ainda que a definição de obrigações aplicáveis a cada caso dependa da redação final da lei (e, em alguns casos, da regulamentação posterior), parece haver um consenso sobre o modelo centrado em riscos ser a própria base sobre a qual o projeto é construído, em alinhamento às referências internacionais. Da mesma forma, o que caracterizará cada nível de risco também poderá ser objeto de alterações e detalhamento posterior, mas há claras tendências internacionais e nacionais (por exemplo, a partir da Lei Geral de Proteção de Dados e posicionamentos da Autoridade de Proteção de Dados Pessoais) de considerar determinadas finalidades ou tratamentos de dados pessoais como de alto risco, a exemplo do auxílio a diagnósticos médicos, recrutamento ou avaliação de empregados ou decisões sobre o acesso a serviços considerados essenciais (incluindo crédito).

O PL brasileiro estabelece obrigações para diferentes agentes da cadeia de IA: desenvolvedores, distribuidores e utilizadores.

A definição mais precisa de obrigações aplicáveis a cada um ainda é nebulosa e, em sua maior parte, somente deve ficar clara após regulamentação específica pelas autoridades competentes – mas o pressuposto de responsabilização e imposição de obrigações legais a toda a cadeia dificilmente será afastado.

As principais obrigações trazidas pelo projeto também estão alinhadas a tendências internacionais e ao que já previa a LGPD:

(i) obrigações de informação e transparência;

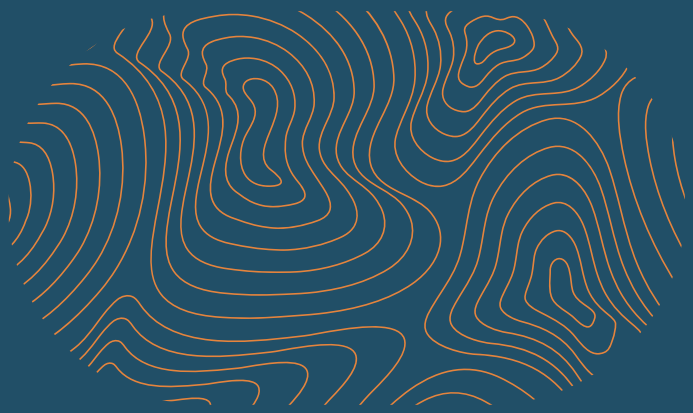
(ii) obrigações de prestação de conta e responsabilização do agente por avaliações de risco e adoção de medidas mitigadoras compatíveis e proporcionais; e

(iii) direitos de revisão e de explicabilidade das decisões tomadas exclusivamente de maneira automatizada (abordadas especificamente no artigo 4 deste e-book).

Novamente, o detalhamento dessas obrigações e como elas serão exigidas dependerá da redação final da lei e de sua posterior regulamentação por autoridades competentes, mas a sua existência decorre de pilares centrais do PL: seus princípios, vários deles em consonância com aqueles também existentes sob a LGPD.

Os últimos 5 anos de debate legislativo culminaram, até o momento, na escolha de um modelo regulatório de **corregulação**, um equilíbrio entre a autorregulação pelo próprio mercado e a imposição rígida de mecanismos de controle pelo Estado. O PL nº 2.338/2023 conta com mecanismos que estimulam a **colaboração** e o **diálogo entre autoridades competentes e agentes econômicos regulados** – um sinal bastante positivo diante de preocupações de que a regulação poderia inibir a inovação. Exemplos destes mecanismos incluem a possibilidade de agentes econômicos criarem seus próprios códigos de conduta, de entidades de mercado tornarem-se certificadoras de boas práticas, além da possibilidade de criação de *sandboxes* regulatórios para promover o aprendizado colaborativo entre autoridades e agentes econômicos.

Outra forte característica do Marco brasileiro é a **abordagem setorial**, com a atuação de diversas **autoridades específicas**, coordenadas pela **Autoridade Nacional de Proteção de Dados (ANPD)**, que também regularia o tema de forma residual.



Aqui, o Brasil parece adaptar novamente o modelo europeu, na qual a regulação também é setorial, mas há certa centralização de autoridades setoriais sob entidades supervisoras "gerais", como o *AI Office* e o *AI Board*. O modelo nacional levou em conta as críticas direcionadas ao AI Act europeu e optou por dar prevalência à regulação setorial, o que deve promover colaboração mais ágil entre autoridades setoriais e os agentes econômicos sujeitos a cada uma delas.

É também por isso que diversos pontos do PL têm o seu detalhamento e regulamentação propositadamente delegados às autoridades competentes – que poderão fazê-lo com maior expertise, de forma setorial e com mais agilidade e flexibilidade do que ocorreria em um processo legislativo.

Ainda não existe lista definitiva das autoridades setoriais que estarão à frente da regulamentação e fiscalização da futura legislação de IA, mas suas competências dificilmente serão uma surpresa: por exemplo, se uma *fintech* já está, atualmente, sujeita ao Banco Central para assuntos regulatórios, inclusive em matérias de proteção de dados pessoais especificamente encontrados no setor financeiro, deve estar atenta, desde já, aos posicionamentos tomados ou sugeridos pelo BACEN também em relação ao uso de IA, mesmo que ainda não exista uma legislação aprovada e em vigor.

COLARES
ADVOGADOS

O avanço do país na regulação de IA segue forte tendência internacional e a crescente expectativa de desenvolvimento e uso responsável de IA por empresas, sobretudo aquelas inseridas em relações econômicas internacionais, não deixa margem para que o assunto continue sem regulação abrangente. O Brasil não está sozinho: dezenas de outros países encontram-se em momento similar, com projetos de legislação específica sob discussão, após a adoção de diretrizes ou políticas nacionais de IA que forneceram as bases gerais para uma futura regulação ser discutida e, eventualmente, aprovada.

Até lá, contudo, **não vivemos um vácuo normativo** — muito pelo contrário. Diferentes legislações nacionais já são aplicáveis ao contexto de IA, ainda que não tenham sido criadas especificamente para tratar do assunto, e diferentes autoridades também já orientam e fiscalizam condutas das empresas sob suas competências.

A LGPD, por exemplo, se aplica a soluções de IA que tratam dados pessoais e prevê obrigações e direitos específicos relativos a tomadas de decisões automatizadas. A ANPD já se debruça sobre assuntos que estão intimamente ligados a soluções de IA e inseriu o tema em sua agenda regulatória do biênio 2025 – 2026. Ou seja, não esperará a tramitação do PL para regular o tema dentro da sua esfera de competência.

Similarmente, o Código Civil, o Código de Defesa do Consumidor (CDC), a Consolidação de Leis Trabalhistas (CLT), a Lei de Direitos Autorais e outras legislações já podem ser (e têm sido) acionadas, a depender do contexto em que a IA seja utilizada, da relação entre a empresa e a pessoa afetada, e da natureza dos impactos que venha a causar. Já se tem casos emblemáticos de ações judiciais relativas ao uso de IA, tanto na Justiça Comum quanto na Justiça do Trabalho, em que diferentes embasamentos legais foram utilizados para responsabilizar empresas que utilizavam IA e escrutinar as medidas adotadas para mitigar os riscos gerados.

Por isso, mesmo com o Marco Legal de IA ainda em discussão no Congresso Nacional, é fortemente recomendável para as empresas que desenvolvem ou utilizam IA adotar, desde já, medidas de governança e boas práticas para mitigar riscos de responsabilização – que efetivamente já existem, especialmente em casos classificados como de alto risco no projeto de lei.

A partir dos moldes gerais do Projeto de Lei nº 2.338/2023, somados aos posicionamentos da ANPD e de autoridades setoriais a respeito da IA, já é possível identificar os primeiros passos necessários à implementação de uma governança interna de IA – e adotar essas medidas desde já não apenas reduz riscos imediatos, como também pode significar um importante diferencial competitivo.

COLARES
ADVOGADOS

O que fazer a partir de hoje?

- Identificar os contextos de usos de IA na empresa
- Avaliar os riscos envolvidos em cada caso
- Criar planos de mitigação para os riscos mais altos, prioritariamente
- Elaborar documentação necessária sob a LGPD (se houver uso de dados pessoais)
- Treinar colaboradores sobre riscos e boas práticas no uso de IA
- Acompanhar posicionamentos da ANPD e de autoridades setoriais às quais a empresa está sujeita
- Ajustar planos de mitigação conforme necessário após aprovação do Marco Legal da IA

Expertise em
Inovação e Tecnologia

COLARES
ADVOGADOS

O **Colares Advogados** é um escritório de advocacia dinâmico, voltado para negócios e disputas complexas, em diversos setores e áreas, com destaque para Contratos, Fusões e Aquisições (M&A), Direito Societário, Inovação e Tecnologia.

Trazemos mais de 20 anos de experiência em assessoria jurídica a empresas intensivas em inovação e tecnologia, em diversos aspectos jurídicos, incluindo no desenho de modelos de negócio inovadores, suas estruturas jurídicas, regulatórias e fiscais, investimentos, contratos e políticas.

Nossa prática em Inovação e Tecnologia inclui a assessoria em aspectos jurídicos e regulatórios relacionados à inteligência artificial, a exemplo da conformidade com frameworks nacionais e internacionais, regulações emergentes e legislações que impactam diretamente o uso da IA, como a Lei Geral de Proteção de Dados e a Lei de Direitos Autorais. Também apoiamos organizações na mitigação de riscos (éticos e legais) em IA, elaboração de contratos e políticas relativos a soluções de IA, e treinamentos corporativos sobre boas práticas no uso de IA.

A atuação do escritório e suas práticas institucionais são reconhecidas pelos principais rankings nacionais e internacionais no segmento jurídico.

Contatos

Débora Vieira
Sócia
debora@colareslaw.com.br
+55 81 99994 4473

Rodrigo Colares
Sócio
rodrigo@colareslaw.com.br
+55 11 99575 9884

COLARES
ADVOGADOS

Recife - São Paulo

WWW.COLARESLAW.COM.BR

Conteúdo, autores e licenciamento

O conteúdo deste e-book foi originalmente produzido pelos seus autores, integrantes do Colares Advogados, com base nas suas experiências profissionais e consulta a informações oriundas de fontes externas. Partes do texto foram revisadas com uso de ferramentas de inteligência artificial generativa, mas todos os dados e suas respectivas fontes foram consultados e incluídos diretamente pelos autores, encontram-se devidamente referenciados e estão disponíveis para acesso pelos hiperlinks inseridos ao longo do texto, que foram verificados pela última vez em 12 de setembro de 2025.

Coordenadora

Débora Vieira

Autores

Débora Vieira
Thaís Praxar
Luana Coimbra
Pedro Leite