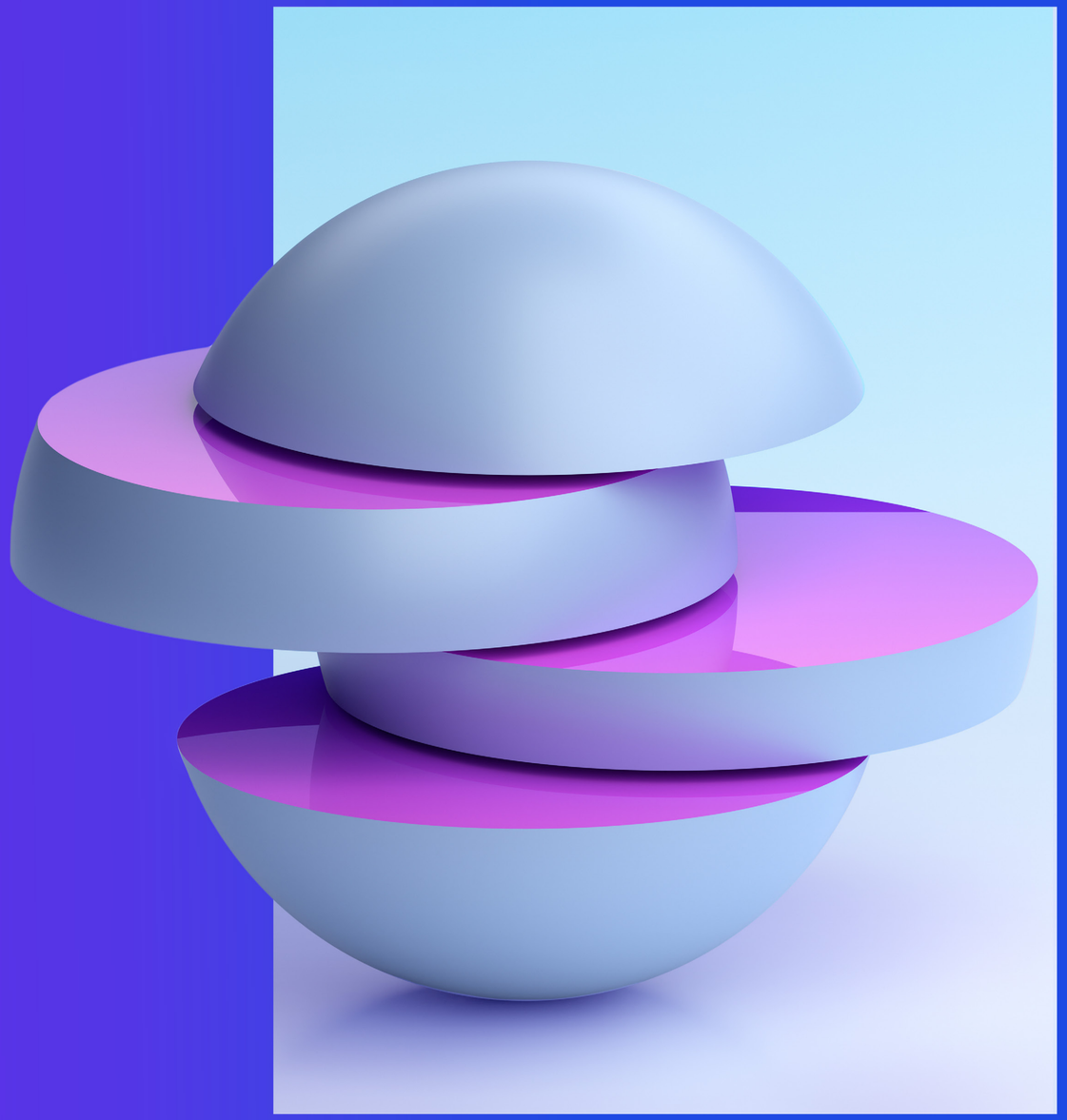
Privacidade em um Novo Mundo de IA

Como construir confiança em IA por meio da privacidade



Dezembro de 2023

kpmg.com.br

Conteúdo

Privacidade de dados e a confiança em IA

A Inteligência Artificial (IA) promete transformar nossas vidas, ajudando-nos a ser mais eficientes, produtivos, saudáveis e inovadores.

Esta empolgante tecnologia já está sendo utilizada em diversos setores e áreas públicas, aproveitando ao máximo o poder dos dados para aprimorar a previsibilidade, melhorar produtos e serviços, reduzir custos e liberar colaboradores de trabalhos administrativos rotineiros.

No setor de serviços de saúde, por exemplo, os médicos podem prever riscos à saúde de forma mais rápida e precisa, além de viabilizarem tratamentos complexos de maneira mais eficaz. Na mineração, robôs movidos a IA estão realizando tarefas perigosas, como mineração de carvão, exploração marítima e também auxiliando em operações de resgate durante desastres. Em serviços bancários e comerciais, a IA e o *Machine Learning* (ML) estão ajudando as equipes de vendas e marketing a identificar potenciais clientes, prospectar e prever suas necessidades e tendências à compra.

Estas tecnologias também permitem a precificação dinâmica dos microssegmentos, bem como automatizam processos de tomada de decisão, conjuntos de regras de crédito e exceções. No setor de consumo e varejo, a IA está ajudando a prever e analisar tendências, criar modelos virtuais que possam exibir roupas, prever as necessidades dos clientes e ajudar os consumidores a desfrutar de uma experiência de compra mais personalizada.

De acordo com o **Global Tech Report 2023 da KPMG**, líderes do setor de tecnologia identificam a IA e o ML como as tecnologias mais importantes para alcançar resultados a curto prazo[1]. Além disso, na pesquisa global ***Trust in Artificial Intelligence*** de 2023, com mais de 17.000 pessoas no mundo inteiro, 85% acreditam que a IA pode proporcionar uma série de benefícios[2].

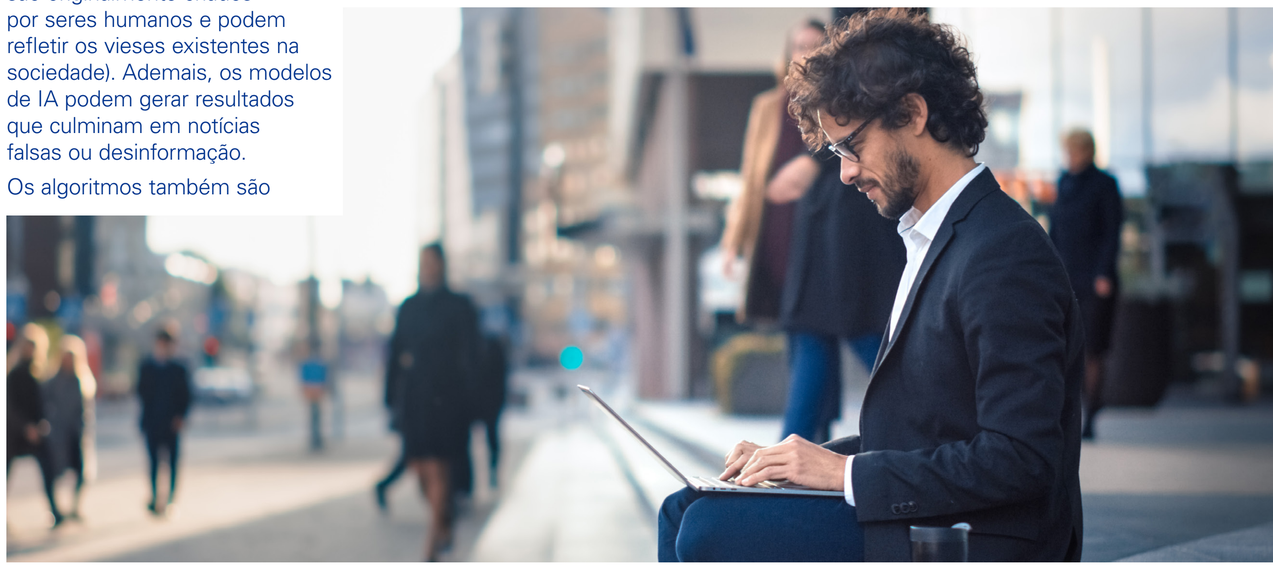
No entanto, assim como qualquer tecnologia emergente, existem riscos. A mesma pesquisa mostra que 61% das pessoas estão cautelosas quanto a confiar nos sistemas de IA, e apenas metade acredita que os benefícios da IA superam os riscos[3]. Adicionalmente, 55% dos líderes de tecnologia dizem que seus avanços com automação estão atrasados por causa de preocupações em relação ao modo como os sistemas de IA tomam decisões[4].O uso generalizado e não regulamentado dessa tecnologia levanta preocupações sobre seu impacto nos direitos humanos e na privacidade. Isso é particularmente verdadeiro para a IA Generativa (GenAI), que é impulsionada por algoritmos baseados em Grandes Modelos de Aprendizagem de Linguagens (LLMs).

No momento da escrita deste artigo, líderes em IA emitiram cartas abertas buscando uma pausa no desenvolvimento da GenAI, incentivando os legisladores a garantir seu futuro por meio de regulamentos, normas e boas práticas[5]. Alguns riscos citados incluem *design* falho, lógica tendenciosa, erros de codificação, vulnerabilidades de segurança e, o mais importante, julgamentos que podem discriminar indivíduos ou grupos (afinal, os dados utilizados são originalmente criados por seres humanos e podem refletir os vieses existentes na sociedade). Ademais, os modelos de IA podem gerar resultados que culminam em notícias falsas ou desinformação.

Os algoritmos também são imprevisíveis, complexos e difíceis de explicar. Devido à sua natureza proprietária, eles carecem de transparência e geram resultados com base no processamento de vastos dados da internet, ampliando o risco de vazamentos de dados confidenciais e violação de dados pessoais legalmente protegidos.

As leis de privacidade internacionais são aplicáveis à coleta de dados em todos os estágios do ciclo de vida da IA, e por isso, não é surpresa que a mineração e coleta de dados tenham resultado em minuciosas verificações de órgãos reguladores globais. Os fiscalizadores das autoridades europeias de privacidade e proteção de dados iniciaram investigações ao redor do mundo sobre a legalidade das atividades de processamento de dados relacionadas à GenAI.

Este artigo investiga as implicações da privacidade na ampla adoção da IA; tem como objetivo entender o que isso significa para as empresas sob a perspectiva da privacidade e delinea as principais medidas que as organizações podem adotar para utilizar a IA de forma responsável. Ao se manterem informadas sobre as implicações da adoção da IA para a privacidade e tomarem medidas proativas para mitigar riscos, as empresas podem aproveitar ao máximo o poder dessa tecnologia enquanto protegem a privacidade dos indivíduos.



[1] KPMG. KPMG Global Tech Report 2023. 2023.
[2] KPMG. *Trust in artificial intelligence.* 2023.
[3] *Ibid.*
[4] KPMG. *Global Tech Report 2023.* 2023.
[5] FUTURE OF LIFE INSTITUTE. *Pause Giant AI Experiments: An Open Letter.* 2023.

O que está por vir? Um cenário regulatório em evolução

À medida que o uso da GenAI aumenta, nações estão se apressando para legislar e criar padrões para o uso responsável da IA.

A União Europeia está realizando a primeira tentativa de legislar o que poderá acabar sendo a legislação de privacidade de IA mais rigorosa em todo o mundo. Além disso, o Conselho Europeu de Proteção de Dados (EDPB) já criou uma força-tarefa dedicada ao ChatGPT para promover a cooperação e o intercâmbio de informações sobre a possível aplicação de leis e normas por parte das autoridades de proteção de dados. No momento da escrita deste artigo, o Conselho e Parlamento Europeu e a Comissão Europeia propuseram a Lei de Inteligência Artificial (AI Act ou Lei de IA), que, quando promulgada, poderá se tornar a legislação global de privacidade de IA mais complexa.

A Lei de IA baseia-se nas disposições de privacidade de acordo com a Regulamentação Geral de Proteção de Dados (General Data Protection Regulation - GDPR), que inclui princípios sobre transparência, equidade, tomadas de decisão algorítmica e dignidade humana. Esses princípios formaram a base dos princípios da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), criados em 2019 para uma gestão responsável da IA, a qual estabelece que os sistemas de IA devem ser robustos, confiáveis e seguros ao longo de seu ciclo de vida.

A Lei de IA reforça ainda mais os princípios da OCDE determinando que a IA seja legal, ética e tecnicamente robusta, e que também respeite os valores democráticos, os direitos humanos e o Estado de Direito. Exatidão, não discriminação, supervisão humana e evidência dos padrões de IA são necessários. A Lei de IA introduz categorias de risco e estabelece requisitos legais para sistemas de IA de alto risco, incluindo aqueles utilizados em avaliações de desempenho, recrutamento e promoções com base no nível de risco público.

A Lei de Serviços Digitais (Digital Services Act - DSA) também foi aprovada pela UE e entrará em vigor em 2024. A DSA impõe transparência, avaliação de riscos e obrigações de responsabilização algorítmica em plataformas de IA, sujeitas a rigorosas auditorias de transparência. O Centro Europeu de Transparência Algorítmica (ECAT) ajudará a reforçar a DSA e tem como objetivo ter relevância

nas novas aplicações de leis e normas digitais no ecossistema de mudanças dentro da UE, e que terá implicações globais, estabelecendo um padrão moral para todo o mundo[6].

Embora os desdobramentos regulatórios em IA também estejam se formando no Canadá, na China e no Brasil em âmbito federal, nos Estados Unidos ainda não surgiu uma abordagem federal para a legislação da IA, apesar de 131 projetos de lei terem sido propostos pelo Congresso dos Estados Unidos. No entanto, regulamentações estaduais e locais para auditorias independentes, modelos de equidade e transparência em relação ao uso de sistemas algorítmicos já foram promulgadas em determinados estados, incluindo Illinois, Maryland, Washington, Nova York e Califórnia.

Os regimes regulatórios que estão atualmente sendo estabelecidos para legislar a IA se sobrepõem significativamente aos regimes regulatórios de privacidade existentes. Essas regulamentações são influenciadas por preocupações globais em relação ao impacto da IA sobre os indivíduos, particularmente em relação à equidade, explicabilidade, transparência, segurança, ao respeito pela pessoa e pela responsabilização. Muitos dos princípios da OCDE sobre IA, desenvolvidos em 2019, podem ser mapeados para princípios de privacidade no contexto de proteção de dados pessoais e estendidos aos Princípios de *Privacy by Design* ao arquitetar os sistemas de IA.

[6] EUROPEAN COMISSION. *Open the black box of algorithms: A deep dive into ECAT's work*. 2023.

Como os princípios de IA da OCDE se relacionam com os princípios de privacidade da GDPR e Privacy by Design.

	Princípios de IA da OCDE				
	• Responsabilização	• Crescimento inclusivo, desenvolvimento sustentável e bem-estar	• Valores humanos e equidade	• Transparência e explicabilidade	• Robustez e segurança
Princípios da GDPR	• Responsabilização	• Limitação de propósito • Minimização de dados	• Legalidade, equidade e transparência • Limitação de propósito • Precisão • Limitação de armazenamento	• Legalidade, equidade e transparência	• Integridade e confidencialidade
Princípios de *Privacy by Design*	• Proativo e não reativo; preventivo e não remediador	• Privacidade incorporada ao projeto • Funcionalidade completa de forma incremental	• Visibilidade e Transparência	• Segurança de ponta a ponta: proteção completa por todo o ciclo de vida	• Segurança ponta a ponta — proteção completa do ciclo de vida

O futuro das regulamentações de IA

O desenvolvimento simultâneo de normas globais de IA e frameworks moldará o futuro das regulamentações de IA. Essas normas e estruturas podem ajudar a proporcionar às organizações uma abordagem mais holística da gestão de riscos de IA e privacidade, especialmente aquelas que estão atrasadas em relação ao mercado.

Por exemplo, o Instituto Nacional de Normas e Tecnologia dos Estados Unidos (US National Institute of Standards and Technology - NIST) introduziu um *framework* de Gerenciamento de Riscos voluntário e autorregulado que visa garantir a credibilidade da IA enquanto mitiga os riscos. A Organização Internacional de Normas (ISO/IEC) também desenvolveu a ISO 42001 em Sistemas de Gestão de Inteligência Artificial (AIMS).

Quando combinadas com os relevantes requisitos regulatórios para os sistemas de IA, essas normas podem servir como ferramentas eficazes de gerenciamento de riscos para ajudar a operacionalizar e promover as melhores práticas e abordagens de governança responsável de IA. Elas estimularão a consistência em diferentes regimes regulatórios, os quais poderão ter requisitos conflitantes em relação à robustez, segurança, equidade e credibilidade dos sistemas de IA.

O atual governo dos Estados Unidos firmou compromissos com grandes empresas de tecnologia para que estas sigam muitos dos princípios e requisitos estabelecidos nessas normas. O tempo dirá como isso se dará em conjunto com as regulamentações globais.

Os principais pontos para alcançar a privacidade em IA

A incorporação de *Privacy by Design* em sistemas de IA deve ajudar a construir confiança e navegar por possíveis desafios de privacidade

A confiança é um fator-chave para a viabilização de receita e crescimento. As organizações que utilizam IA devem incorporar privacidade nos processos de desenvolvimento de sistemas para garantir que estes sejam seguros, eficazes e imparciais, e sejam respaldados por uma governança firme, responsabilização clara e supervisão robusta. Enquanto aguardam as legislações se atualizarem em relação às tecnologias, as organizações que desejam lançar sistemas em IA devem integrar a privacidade em cada etapa do ciclo de vida como uma das melhores práticas. A adoção da abordagem *Privacy by Design* pode ajudar a assegurar aos clientes, órgãos reguladores e demais partes interessadas sobre a credibilidade da IA e minimizar quaisquer impactos negativos.

Visando auxiliar as organizações a adotar uma abordagem proativa de engenharia de *Privacy by Design*, apresentamos a seguir os principais princípios de privacidade que devem ser considerados ao longo de todo o ciclo de vida da IA.

1 Legalidade e equidade

A IA possui um propósito legítimo, legal e claramente definido, exercendo um impacto mínimo sobre a privacidade.

Risco à privacidade

Falhas na concepção e na implementação de modelos, bem como ausência de mecanismos de segurança adequados.

Tais riscos podem surgir quando as especificações projetadas não atendem adequadamente às tarefas desejadas, possivelmente devido a inadequações nos dados de entrada, nas variáveis escolhidas, em premissas falhas ou na priorização excessiva da eficiência em detrimento da eficácia.

O que pode dar errado

Uma agência do setor público implementou um sistema guiado por IA para detectar fraudes em benefícios. No entanto, os resultados obtidos eram imprecisos, aleatórios e injustos. Dentre as falhas centrais na concepção e na execução do modelo estava a inclusão de filtros de inputs invasivos, irrelevantes, banais e subjetivos. Por exemplo, o sistema considerou status de relacionamento, duração e frequência para atribuir pontuações de alto risco a certos candidatos, ignorando, porém, outros. Além disso, o modelo estabeleceu correlações sem fundamento entre diferentes entradas, comprometendo a consistência das decisões. E, para agravar, o mecanismo de tomada de decisão carecia de uma via eficaz para contestações e análises de decisões, um componente essencial para garantir a segurança.
Como consequência, o algoritmo, sem apresentar evidências claras ou justificativas plausíveis, classificou de forma desproporcional pessoas que realmente necessitavam dos benefícios como de alto risco de fraude. Candidatos elegíveis, identificados como alto risco, enfrentaram investigações invasivas e foram estigmatizados. Assim, houve atrasos indevidos no recebimento de benefícios tão essenciais para eles.

Mitigação do risco

É fundamental selecionar os inputs com critério, assegurando que sejam relevantes, legais e não discriminatórios. Antes de tomar decisões alicerçadas em correlações, é vital garantir uma relação clara de causa e efeito..

2

Transparência e explicabilidade

A transparência é essencial para a responsabilidade e para a otimização do produto. A explicabilidade e a interpretabilidade facilitam a compreensão nas tomadas de decisão, proporcionando confiança quando a IA opera de maneira eficaz e possibilitando revisões quando falha.

Risco à privacidade

Não conformidade com a regulamentação; ameaças à propriedade intelectual; ausência de mecanismos de segurança; falhas na construção do modelo

O que pode dar errado

A incapacidade de explicar e justificar as decisões advindas de sistemas de IA lastreados por algoritmos do tipo "caixa preta" impede que os indivíduos questionem o processo e seus resultados. A ausência de transparência, tanto nas informações fornecidas quanto nos resultados gerados, dificulta a prevenção de desfechos discriminatórios ou nocivos.

Mitigação do risco

As informações pessoais usadas para treinar conjuntos de dados devem ser transparentes, confiáveis, precisas, íntegras e pertinentes aos resultados gerados, e devem poder ser questionadas caso mostrem-se tendenciosas ou imprecisas. Todos os dados coletados precisam estar acompanhados de avisos sobre privacidade (direito de ser informado)..

3

Governança e responsabilização

As leis de privacidade demandam estruturas de governança sólidas e programas de privacidade bem estruturados, que estabeleçam e comuniquem de forma nítida suas funções e responsabilidades.

Risco à privacidade

Alteração de propósito; uso de dados de maneira incompatível com sua finalidade original; projeto inadequado de algoritmo.

Frequentemente, os desenvolvedores de IA ou os gestores de projetos tendem a superestimar ou até mesmo fazer alegações infundadas sobre seus modelos. Quando os compradores confiam nessas afirmações sem realizar os devidos testes e homologações, as insuficiências só se manifestam após danos reais ou outras falhas associadas à IA. Por exemplo, "falsos positivos" em reconhecimento facial podem levar a detenções indevidas, aprisionando inocentes e invadindo injustificadamente a privacidade das pessoas

O que pode dar errado

Um barco autônomo criado para traçar a rota mais rápida através de um porto pode, inadvertidamente, prejudicar ecossistemas aquáticos sensíveis e surpreender nadadores, caso não seja programado para navegar com segurança. Assim, os dados de treinamento para tal sistema devem incluir imagens de nadadores.

Mitigação do risco

Sistemas de IA necessitam de clareza nas linhas de responsabilidade para assegurar que: os riscos associados à IA sejam geridos eficientemente; o uso da IA tenha um propósito definido, uma estratégia e um conjunto claro de expectativas comunicadas a todas as partes envolvidas; exista supervisão adequada, treinamento e transparência nas informações, sobretudo no que diz respeito a limitações e dados sensíveis; terceiros, como fornecedores de serviços de dados ou desenvolvedores de IA, sejam igualmente responsabilizados e colaborativos na resolução de eventuais problemas que possam emergir.

4

Minimização de dados

Os dados pessoais devem ser adequados, relevantes e limitados ao seu propósito.

Risco à privacidade

Falhas pós-implementação, como problemas de robustez, ataques externos e interações imprevistas.

No universo da IA, prevalece a ideia de que "quanto mais, melhor". No entanto, nem todos os dados possuem a mesma qualidade. Sistemas treinados com conjuntos de dados que não passaram por validação externa podem não apresentar o desempenho esperado no mundo real. Isso pode levar ao uso impróprio e à utilização injusta das informações pessoais. Além disso, falantes de línguas sub-representadas muitas vezes não são consultados antes que seus idiomas sejam usados para treinar e desenvolver modelos de linguagem natural, o que pode negligenciar aspectos culturais relevantes.

O que pode dar errado

Conjuntos de dados que não passaram por validação externa podem gerar recomendações ou interações inadequadas com menores, propagação de desinformação, manifestações preconceituosas, respostas logicamente incoerentes ou falsas. Além disso, o uso de imagens de baixa resolução no treinamento de sistemas de reconhecimento facial pode comprometer a precisão em aplicações práticas e introduzir vieses.

Mitigação do risco

Não são todos os dados que se mostram úteis, relevantes ou confiáveis. A prática de minimização de dados estimula uma seleção mais criteriosa para os treinamentos. Ao excluir dados inapropriados, o conjunto de dados resultante, embora menor, é de maior qualidade. O treinamento de LLMs predominantemente em inglês ou em outros idiomas majoritários pode gerar modelos enviesados, excluindo diversas línguas do conjunto de dados, especialmente se forem consideradas apenas certas fontes, como internet e redes sociais.

5

Limitação de propósito

O processamento de dados pessoais deve ter um propósito claramente definido e ser devidamente comunicado, a fim de proteger os direitos do titular, garantir sua autonomia e prevenir possíveis danos.

Risco à privacidade

Falhas na elaboração e implementação de modelos, incluindo mudança de conceito e desalinhamento de dados; tarefas não realizáveis; não conformidade com regulamentações; modelos de negócio ilegítimos.

A ausência de um propósito claro pode levar a informações distorcidas e tendenciosas. Por exemplo, ao utilizar o trabalho escrito de um autor para produzir novos artigos e, falsamente, atribuir a autoria a ele.

O que pode dar errado

Uma LLM (inserir termo quando apresentado pela primeira vez) treinada com base em artigos de notícias públicos e outros dados colhidos da internet pode, inadvertidamente, propagar informações incorretas ou disseminar fake news e desinformações. Se forem produzidos artigos difamatórios, prejudiciais e polêmicos, imitando o estilo de um autor cujo material serviu de base para o treinamento do modelo de IA, isso pode afetar adversamente a reputação do autor real, caso as pessoas acreditem que foi ele quem realmente escreveu.

Mitigação do risco

Todas as partes envolvidas com sistemas de IA, desde desenvolvedores até vendedores e usuários finais, devem compreender e honrar o propósito estabelecido. Isso guiará a seleção de dados usados no treinamento do modelo, as situações de uso durante a implementação e a operação, os princípios e suposições integrados ao sistema, sua configuração, suas medidas de proteção e muito mais. Agindo assim, as decisões iniciais relacionadas ao consentimento e suas exceções são observadas, fortalecendo a confiança do público e diminuindo o risco de descumprimento regulatório e de reações adversas por parte da sociedade.

6

Precisão

Conforme as leis de privacidade, antes de serem utilizados, os dados pessoais devem ser atualizados, completos e precisos. Adicionalmente, os indivíduos têm o direito de corrigir seus dados sempre que necessário.

Risco à privacidade

Questões de robustez, seja pela subutilização ou superutilização de recursos; vulnerabilidades a ataques externos; interações imprevistas; ausência de mecanismos de segurança; práticas injustas; alteração no modelo.

A integridade dos dados é fundamental para o desempenho eficiente da IA, e sua negligência pode conduzir a múltiplos prejuízos. Por exemplo, informações imprecisas podem afetar decisões em políticas governamentais ou planejamentos comunitários. Ainda que os dados de entrada estejam corretos, falhas no modelo podem gerar uma representação inadequada do indivíduo devido a suposições incorretas, classificações errôneas ou a incapacidade de lidar com informações desconhecidas. Dados pessoais desatualizados ou não pertinentes podem gerar viéses no modelo e comprometer seu desempenho, como, por exemplo, reduzindo a acuracidade de previsões. Sem uma revisão humana adequada, um sistema de IA pode produzir decisões errôneas, mesmo operando conforme programado e sem identificar irregularidades.

O que pode dar errado

Suponha que a polícia detenha equivocadamente um indivíduo porque seu sistema de reconhecimento facial o apontou como suspeito, mesmo sem mais provas. Mesmo com dúvidas sobre a similaridade entre o indivíduo e a foto, as autoridades confiam nas alegações de precisão fornecidas pelo fabricante do sistema de IA. Como resultado, esse indivíduo enfrenta prejuízos financeiros por salários não recebidos e despesas jurídicas, além de humilhação, ansiedade, inconvenientes, perda de liberdade e possíveis traumas decorrentes da detenção indevida.

Mitigação do risco

Os princípios de minimização e acuracidade dos dados podem ser instrumentais para elevar a qualidade das informações e, assim, evitar muitos desses problemas. A título de exemplo, o Information Commissioners Office (ICO) do Reino Unido orienta os desenvolvedores de IA a avaliar o equilíbrio entre dados minimizados e precisão estatística durante a fase de testes, assegurando a acuracidade do modelo[7]. Torna-se imperativo monitorar a performance do modelo, especialmente em contextos decisórios ou em sistemas de IA preditivos, para certificar que os dados se mantêm pertinentes, atualizados e corretos, e que sejam reajustados quando necessário.

[7] UK INFORMATION COMMISSIONERS OFFICE. *Guidance on AI and data protection. 2023.*

7

Limitação de armazenamento

As leis de privacidade exigem que as empresas não retenham dados pessoais além do tempo necessário para seu uso. Contudo, em certas situações, é permitida a retenção desses dados, desde que sejam devidamente anonimizados.

Risco à privacidade

Falta de conformidade com a regulamentação;
Após o treinamento do modelo, os dados subjacentes devem ser mantidos apenas se houver necessidade de um novo treinamento. Ainda assim, devido ao risco de desvio do modelo, é crucial realizar uma reavaliação periódica da qualidade dos dados para eliminar informações obsoletas ou irrelevantes.

Manter conjuntos de dados de treinamento após seu uso, mesmo quando anonimizados, acarreta desafios significativos em conformidade, pois o risco de reidentificação é constante e demanda monitoramento contínuo.

Os órgãos reguladores estão especialmente atentos aos riscos de vazamentos, reutilização indevida, e enriquecimentos ilícitos de dados, além do perigo de reidentificação associado à manutenção de dados por períodos prolongados. Além disso, dados de treinamento que vazem, mesmo que anonimizados, podem ser cruzados com outras fontes ou submetidos a engenharia reversa para reidentificar os indivíduos envolvidos.

O que pode dar errado

Considere um modelo de IA desenvolvido para avaliar a adequação de um candidato a uma vaga de emprego, que foi treinado com um conjunto de dados antigo, priorizando a capacidade do candidato de estar fisicamente presente no local de trabalho. Se esse modelo não for atualizado para considerar a nova realidade do trabalho remoto, candidatos que optem por essa modalidade podem ser injustamente desfavorecidos, sendo excluídos ou recebendo uma pontuação menor no processo seletivo.

Mitigação do risco

É fundamental estar a par de todas as leis de retenção de dados, revisando-os e atualizando-os, quando necessário, de forma contínua.

8

Segurança

As empresas responsáveis pelo processamento de dados pessoais devem assegurar sua confidencialidade, integridade e disponibilidade..

Risco à privacidade

Vulnerabilidade a ataques externos; não conformidade com regulamentações; perda de confiança.

Práticas de segurança inadequadas podem resultar na violação dos dados de treinamento, os quais podem conter informações sensíveis como detalhes financeiros, dados demográficos e códigos postais. Uma tal violação poderia expor indivíduos contidos no conjunto de dados de treinamento a riscos como fraude de identidade, prejuízos financeiros, ansiedade e transtornos ao tentar prevenir danos subsequentes.

O que pode dar errado

Entre os principais riscos à segurança, destacam-se::

- Reidentificação através de ataques do tipo “caixa preta” e “caixa branca”.
- Potencial de divulgação/vazamento, com o risco de inferir informações a partir de dados supostamente anônimos.
- • Violações de dados por meio de ataques externos, como acessos não autorizados aos sistemas.

Mitigação do risco

Uma análise profunda de segurança exigirá expertise mais especializada. Contudo, certos aspectos da privacidade em segurança cibernética, como os mencionados acima, necessitam de atenção particular, uma vez que impactam diretamente nos requisitos de confidencialidade, integridade e disponibilidade..

9

Respeito à privacidade do usuário final

A IA deve honrar os direitos de privacidade, abrangendo direitos como acesso às informações, correção, explicação, exclusão e decisão automatizada.

Risco à privacidade

Não conformidade com a regulamentação; perda de confiança; desafios nas relações públicas; ausência de mecanismos de segurança.

O que pode dar errado

Imagine um sistema de IA que negligencia os direitos de privacidade e cujas verificações são inadequadas. Isso colocaria a empresa em situação de vulnerabilidade, expondo-a a riscos tanto regulatórios quanto reputacionais.

Mitigação do risco

Estes direitos são pertinentes durante todo o ciclo de vida da IA, embora possam ter nuances em diferentes etapas. Mecanismos de explicabilidade devem ser integrados em todas as fases da IA, uma vez que o design, os critérios de entrada e a modelagem têm o potencial de influenciar as decisões tomadas. Assegurar a possibilidade de revisar e questionar estes direitos de privacidade garante que os indivíduos mantenham o controle sobre seus dados pessoais. O direito de corrigir dados inseridos ou de desafiar premissas baseadas em um determinado modelo é crucial para garantir decisões justas e acuradas.

Destaques regulatórios*

Os países estão focando em supervisionar a rápida adoção da IA e garantir que o futuro contemple tanto benefícios para os negócios quanto salvaguardas públicas claras para privacidade e confiança.

A **Austrália** divulgou oito princípios voluntários de ética em inteligência artificial, concebidos para assegurar o uso da IA de maneira segura e confiável.

O **Brasil** definiu princípios, normas e diretrizes para regular o desenvolvimento e a aplicação da IA. Está na vanguarda da formulação de políticas de IA na região e, dada sua magnitude e importância na América Latina, a *estrutura de IA* proposta pelo Brasil tem o potencial de estabelecer tendências em toda a região.

O **Canadá** apresentou o Projeto de Lei C-27, que engloba a Lei de Inteligência Artificial e de Dados, com o objetivo de estabelecer novas normas para o desenvolvimento e a aplicação responsável de IA. O governo federal divulgou, ainda, um *documento complementar* para enriquecer a estrutura sugerida pela Lei. Esta iniciativa representa o primeiro passo em direção a um novo marco regulatório, concebido para direcionar positivamente a inovação em IA e promover a adoção responsável destas tecnologias por empresas e cidadãos canadenses.

A **China**, por sua vez, emitiu *diretrizes de gestão para recomendações algorítmicas em serviços de informação na internet*, visando proteger a segurança nacional e os direitos e interesses de seus cidadãos. Adicionalmente, a China sugeriu medidas para a gestão de serviços impulsionados por inteligência artificial.

A **França** apresentou uma estratégia nacional para a IA, centrada em três metas principais: atingir a excelência científica em IA, fortalecendo e atraindo os melhores talentos globais; estimular o investimento em IA; e assegurar uma abordagem ética para seu uso, mantendo a proteção da privacidade. A Comissão Nacional de Informática e Liberdades (CNIL) divulgou uma página dedicada a recursos sobre IA, que traz um guia detalhado para auto avaliação e planos de ação preparatórios para a Lei de IA da UE, abordando ainda inovações recentes no campo da IA, como o GenAI.

Em 2018, a força-tarefa sobre inteligência artificial da **Índia** ofereceu sugestões de políticas destinadas à aplicação ética da IA, propostas para orientar o governo durante um período de cinco anos.

O **Japão** estabeleceu princípios voltados à promoção de uma IA que respeite os aspectos humanos, visando assegurar a privacidade e segurança dos dados e fomentar um ambiente em que a sociedade possa se beneficiar das informações fornecidas pelos indivíduos às instituições. Ademais, o Ministério da Economia, Comércio e Indústria divulgou diretrizes sobre a governança na aplicação e nos princípios da IA. Estes *frameworks* em desenvolvimento buscam assegurar que o Japão evolua rumo a uma sociedade adaptada à IA.

A **Nova Zelândia** divulgou diretrizes éticas baseadas no Tratado de Waitangi/Te Tiriti e nos princípios Māori para IA, algoritmos, dados e IoT (inserir termo quando apresentado pela primeira vez). Estas orientações são dirigidas a entidades governamentais e outras que lidam com os dados das comunidades Māori.

Em 2020, a **Arábia Saudita** apresentou sua estratégia nacional para dados e IA. Além disso, em 2022, sua Lei de Proteção de Dados Pessoais foi revisada para assegurar uma proteção abrangente de dados durante o uso de informações ligadas à privacidade individual.

Singapura estabeleceu uma estrutura voluntária de Verificação de IA, incentivando empresas a serem mais transparentes sobre as capacidades de seus sistemas de IA, visando a manter as partes interessadas bem informadas e a fomentar confiança na IA. A Comissão de Proteção de Dados Pessoais (PDPC) lançou seu *framework* de governança para modelos de IA, complementado por uma série de recursos, entre eles um guia de implementação e autoavaliação para organizações, casos de uso, informações sobre a Estrutura de Testes para Verificação de IA de Singapura e ferramentas de software. Adicionalmente, a Autoridade Monetária desse País divulgou seus Princípios para promover a equidade, ética, responsabilidade e transparência (FEAT) na aplicação de inteligência artificial e análise de dados no setor financeiro do país.

A **Espanha** anunciou a criação da primeira agência europeia dedicada à supervisão de IA: a Agência Espanhola de Supervisão da Inteligência Artificial (AESIA). Com esta iniciativa, a Espanha almeja ser pioneira na regulamentação da IA na Europa. Recentemente, a Autoridade Espanhola de Proteção de Dados (AEPD) também forneceu diretrizes para a auditoria de atividades relacionadas ao processamento de dados que envolvem IA.

A Comissão de Proteção de Dados Pessoais da Coreia do Sul elaborou um *checklist* para mapeamento e tratamento de dados em processos utilizando IA.

Os **Emirados Árabes Unidos (UAE)** desenvolveram um conjunto de ferramentas éticas para IA com o objetivo de orientar a indústria e o público sobre o uso responsável de sistemas de IA, especialmente diante da ausência de legislação específica para regular essa tecnologia. Em outubro de 2017, o governo dos Emirados Árabes Unidos apresentou sua estratégia oficial para inteligência artificial.

O **Reino Unido** estabeleceu como prioridade avançar na adoção adequada da IA, reconhecendo seu potencial como um "risco elevado" para os direitos e liberdades individuais. É enfatizado que a confiança do público é crucial para uma implantação segura da IA. A Estratégia Nacional de IA traça um ambicioso plano de 10 anos com o intuito de manter o Reino Unido na vanguarda global da IA. A abordagem proposta pelo RU (indicar termo quando apresentado plea primeira vez) favorece a inovação na regulamentação da IA. O Gabinete do Comissário da Informação (Information Commissioner's Office - ICO) disponibilizou recursos sobre IA e proteção de dados, incluindo o Conjunto de Ferramentas para IA e Proteção de Dados, para auxiliar na concretização desse plano. O Centro de Ética e Inovação em Dados (CDEI) também divulgou seu "Barômetro para IA".

Os **Estados Unidos**, por sua vez, ainda não apresentaram uma legislação ou regulamentações federais específicas para IA. Contudo, vários estados já introduziram leis relacionadas à IA em âmbito estadual, enquanto outros estão trabalhando para fazer o mesmo. O Instituto Nacional de Normas e Tecnologia (NIST) elaborou um *framework* para gestão de riscos associados à IA, e a Comissão Federal de Comércio (FTC) publicou orientações relativas ao uso de inteligência artificial e algoritmos.

* Estes destaques não são uma lista exaustiva, mas apenas ilustram o cenário atual no momento da redação deste estudo.

O caminho a seguir: construindo uma IA confiável

A privacidade de dados é o alicerce de uma empresa que presta serviços de IA, precisando ser transparente sobre como as informações pessoais são utilizadas, e assumindo total responsabilização por qualquer uso indevido, além de possuir processos para uma rápida mitigação de incidentes, caso necessário.

A velocidade e a eficiência da IA estão transformando o mundo, e as empresas estão compreensivelmente interessadas em aproveitar ao máximo seu potencial. A vantagem tecnológica só irá oferecer benefícios de mercado se os clientes e outras partes interessadas confiarem que os dados estão sendo utilizados com responsabilidade.

Embora os órgãos reguladores ainda estejam enfrentando dificuldades para acompanhar o ritmo dos avanços na IA, a Lei de IA da União Europeia transmite uma forte mensagem de que a regulamentação será abrangente e trará consequências significativas para quem não estiver em conformidade. As empresas que desenvolvem controles robustos sobre o uso da IA com base em claras diretrizes éticas devem estar em uma posição para alavancar os benefícios da IA enquanto protegem a sociedade contra potenciais riscos e também satisfazem os requisitos regulatórios

Cinco passos essenciais que podem ajudar as empresas a desenvolverem a confiança na IA

1 Entenda o seu contexto regulatório e adote uma estratégia de privacidade de IA

Os legisladores, os responsáveis pela elaboração de políticas e os órgãos reguladores enfatizam consistentemente o alinhamento dos sistemas de IA com as normas reconhecidas. Dessa forma, é essencial identificar quais estruturas regulatórias se aplicam à sua empresa, determinando com quais legislações você deve cumprir e então planejar como a sua IA será implementada. Crie uma base de referência para o uso da IA que satisfaça diferentes metodologias e simplifique o seu desenvolvimento de IA ou suas atividades de negócio relacionadas à IA.

2 Incorpore *Privacy by Design* aos projetos de IA

Avalie o impacto sobre a privacidade e aborde as questões de compliance no estágio de planejamento, e ao longo de todo o ciclo de vida da IA por meio de uma avaliação do impacto sobre a privacidade (PIA) ou da avaliação do impacto na proteção de dados (DPIA). Controles de Privacy by Design, conforme elaborado na norma ISO 31700 de Privacy by Design e no framework de Avaliação de Privacidade no Conceito da KMPG, pode ajudar as organizações a incorporar a privacidade nos sistemas de IA.

Mesmo que você acredite que o seu sistema só utilize dados anônimos ou não pessoais, os riscos à privacidade podem surgir, incluindo a reidentificação de conjuntos de dados de treinamento e até de modelos de IA, além de impactos derivados por utilizar dados não pessoais para treinar modelos que possam impactar indivíduos e comunidades. Uma avaliação robusta também pode incluir uma modelagem das ameaças à segurança e à privacidade ao longo do ciclo de vida da IA, consultando as partes interessadas quando necessário. Considere questões de privacidade mais amplas, tais como a justiça de dados (se as pessoas são adequadamente tratadas de acordo com a maneira que seus dados são utilizados) e a soberania de dados dos povos indígenas (os direitos dos povos indígenas de governar dados sobre suas comunidades, seus povos, suas terras e seus recursos).

3 Avalie os riscos à privacidade de IA

Sempre avalie os riscos à privacidade pela IA, seja em conjunto com o seu PIA/DPIA, ou até depois. Certifique-se de aderir às principais normas, melhores práticas e aos códigos de conduta para operacionalizar os requisitos. Isso pode ser aplicável se você é o desenvolvedor ou o cliente que está adquirindo e integrando um sistema de IA. Se você for um cliente, peça ao desenvolvedor que lhe forneça a documentação essencial e preste suporte às suas avaliações de risco de privacidade de IA e ao PIA. Caso eles não possam fornecer essa documentação, considere outro prestador. Em muitas jurisdições, incluindo o Reino Unido e da UE, o PIA/DPIA é uma exigência legal.

4 Audite seu sistema de IA

Seja você um desenvolvedor, intermediário ou fornecedor de sistemas de IA; é fundamental assegurar aos seus clientes e órgãos reguladores que todos os cuidados foram tomados na elaboração de uma IA confiável. Uma estratégia é realizar auditorias internas em conformidade com normas, regulamentações e melhores práticas do mercado, avaliando também o impacto algorítmico nos sistemas.

5 Respeite os direitos e escolhas por meio da explicabilidade e transparência sobre as entradas e saídas de dados

Esteja pronto para esclarecer dúvidas dos titulares de dados e administrar suas preferências diante do desenvolvimento ou aplicação de sistemas de IA. As organizações que buscam adotar a IA para decisões automatizadas devem ser capazes de explicar de forma clara e objetiva como isso impactará os usuários finais. A "explicabilidade" refere-se à capacidade de elucidar por que a IA tomou determinada decisão ou fez certas previsões. É aconselhável documentar os fluxos de trabalho para identificar e detalhar quais dados foram usados, como foram aplicados ao usuário final e como os titulares podem contestar decisões tomadas pela IA.

Como isso se conecta com o que fazemos.

A KPMG está orientando empresas de diversos setores rumo a uma nova era de oportunidades na economia digital. Da concepção à execução, os profissionais da KPMG estão prontos para auxiliar na transformação de seu modelo de negócios, potencializando competitividade, crescimento e valor futuro.

Com uma trajetória que se estende por mais de 150 anos, a KPMG tem desempenhado um papel crucial na exploração e utilização de novas tecnologias, como a GenAI, além de fornecer consultoria e garantias sobre estas tecnologias.

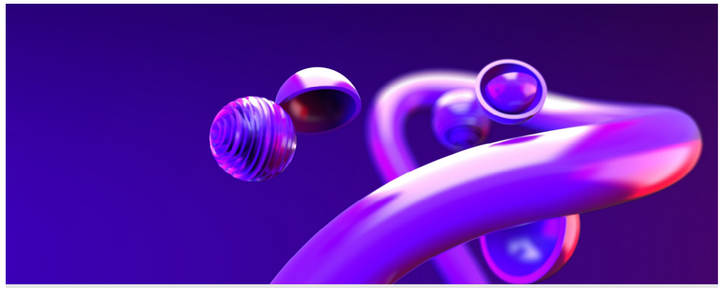
Reconhecemos a complexidade do desafio que é desenvolver uma IA responsável, envolvendo aspectos corporativos, regulatórios e técnicos. A KPMG está dedicada a auxiliar os clientes a entregar serviços de IA de maneira consciente. Através do uso responsável da GenAI, a KPMG auxilia organizações a criar soluções de tecnologia de IA que sejam confiáveis e seguras.

Ademais, nossos especialistas em gestão de riscos de privacidade adotam uma postura responsável ao avaliar a ética, governança e segurança empregadas nas tecnologias de IA e Machine Learning de nossos clientes.

Nosso propósito é fornecer uma visão abrangente sobre tendências de privacidade e disponibilizar um leque diversificado de soluções e serviços, com destaque em PrivaTech, *Privacy by Design*, modelos operacionais voltados para privacidade, ESG, design de programas de privacidade, implementação, automação, experiência do cliente, gestão de mudanças e respostas a quebras de privacidade.

KPMG Connected Enterprise

A abordagem ágil e centrada no cliente para a transformação digital, adaptada para cada setor.



KPMG Powered Enterprise

O conjunto de serviços da KPMG para transformar funções. Modelos operacionais projetados para o futuro, usando as práticas e processos da KPMG e plataformas SaaS (Software as a Service) pré-configuradas.

KPMG Trusted

Como construir e manter a confiança dos *stakeholders*.

KPMG Elevate

Viabilize a geração de valor com rapidez e confiança.

Fale com o nosso time

Leandro Augusto
Sócio-líder de Cyber Security & Privacy da KPMG no Brasil e na América do Sul
lantonio@kpmg.com.br

Ricardo Santana
Sócio-líder de Data & Analytics da KPMG no Brasil
santana@kpmg.com.br

Alguns ou todos os serviços aqui descritos podem não ser permitidos para clientes de auditoria da KPMG e suas afiliadas ou entidades relacionadas.

kpmg.com.br

/kpmgbrasil