Contribuições da Brasscom à Tomada de Subsídios ANPD sobre Inteligência Artificial e Revisão de Decisões Automatizadas

INTRODUÇÃO

A Brasscom agradece a oportunidade de apresentar comentários à Tomada de Subsídios da ANPD sobre Inteligência Artificial e Revisão de Decisões Automatizadas, e parabeniza a Autoridade pela iniciativa.

Antes de adentrarmos as respostas da tomada de subsídios, destacamos que é fundamental que se alinhe e conceitue muito bem do que está sendo discutido e, no futuro regulado. A distinção entre sistemas complexos, sistemas de IA e modelos de IA é um exemplo fundamental de como conceitos técnicos distintos precisam ser tratados como tal, sob pena de gerar impactos indesejados em decisões regulatórias.

De acordo com a ISO/IEC 22989:2022[1], um modelo de IA é definido como uma "representação física, matemática ou lógica de um sistema, entidade, fenômeno, processo ou dados". Em termos práticos, refere-se ao elemento central que trabalha com entradas para produzir saídas, como recomendações e inferências. Esses modelos podem incluir funções matemáticas, redes neurais e árvores de decisão, podendo ser desenvolvidos manualmente ou por técnicas de aprendizado de máquina. Contudo, os modelos, isoladamente, não possuem funcionalidade prática sem o suporte de outros elementos[2].

Por outro lado, ainda de acordo com a ISO/IEC 22989:2022, um sistema de IA é definido como um "sistema projetado para gerar saídas, como conteúdo, previsões, recomendações ou decisões, alinhadas a objetivos definidos por humanos". Um sistema de IA é um conceito mais abrangente, que engloba o modelo e outros componentes, como interfaces, sensores e softwares convencionais.

De acordo com a OCDE[3], a construção de sistemas de IA envolve a integração de um ou mais modelos com outros elementos, como interfaces de usuário e software adicional, para criar um sistema funcional. Ou seja, modelos são um componente dentro de uma cadeia de valor mais ampla. Adicionalmente, um sistema de IA pode ser parte de um

[1] ISO/IEC 22989:2022. Information technology — Artificial intelligence — Artificial intelligence concepts and terminology. Edition 1, 2022. International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC).

[2] Fernández-Llorca, D., Gómez, E., Sánchez, I. et al. An interdisciplinary account of the terminological choices by EU policymakers ahead of the final agreement on the AI Act: AI system, general purpose AI system, foundation model, and generative AI. Artif Intell Law (2024). https://doi.org/10.1007/s10506-024-09412-y

[3] OECD (2024), "Explanatory memorandum on the updated OECD definition of an AI system", OECD Artificial Intelligence Papers, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en

sistema complexo ainda maior. Usando o exemplo de armazéns conectados, o modelo de IA é responsável por dotar o estabelecimento de visão computacional, processando dados para facilitar o armazenamento de produtos e o gerenciamento de estoques. Esse modelo é integrado a um sistema de IA mais amplo, que combina a visão computacional com mecanismos para auxiliar na logística de distribuição de mercadorias para os clientes. Finalmente, o armazém conectado como um todo é um sistema complexo que integra o sistema de IA, além de outros subsistemas, para otimizar a identificação da mercadoria, através da leitura do código, o empacotamento do objeto e até mesmo a impressão da etiqueta com informações da entrega.

Essas distinções são cruciais em contextos regulatórios. Isso porque, ainda que todos esses elementos possam envolver tratamento de dados pessoais, é essencial destacar que, quando se fala em treinamento, o foco está nos modelos, e não nos sistemas complexos.

Segue abaixo nossas respostas às perguntas disponibilizadas:

BLOCO 1 – PRINCÍPIOS DA LGPD

1.Como compatibilizar o treinamento de sistemas de IA com o princípio da necessidade, haja vista se tratar de atividade que, muitas vezes, demanda o tratamento de quantidades massivas de dados pessoais? Quais salvaguardas podem ser adotadas de modo a assegurar a observância desse princípio e viabilizar o desenvolvimento adequado de sistemas de IA, considerando, ainda, a importância da qualidade e diversidade dos dados utilizados?

RESPOSTA: A IA é um conjunto de tecnologias capaz de aprender, adaptar-se e realizar tarefas de maneiras similar a humana. Com acesso a dados, poder computacional e engenhosidade humana necessários para extrair cada vez mais valor deles, os pesquisadores estão criando softwares e máquinas inteligentes para aumentar a produtividade humana e capacitar as pessoas em todo o mundo. Já estamos vivenciando como a IA beneficia as pessoas, a sociedade e a economia em uma ampla gama de áreas. Suas aplicações abrangem diversos campos.

A eficiência e a eficácia dos sistemas de IA estão intimamente relacionadas à qualidade dos dados usados durante o treinamento. Embora o treinamento de sistemas de IA frequentemente envolva o tratamento de grandes quantidades de dados, é importante destacar que esses dados não são necessariamente pessoais. Com frequência são utilizados dados estatísticos e técnicas de anonimização, limitando ou extinguindo riscos no que tange a dados pessoais.

Embora dados pessoais possam ser utilizados no treinamento da IA, eles geralmente não são a principal fonte para o desenvolvimento. De fato, dados não pessoais

frequentemente desempenham um papel mais significativo e relevante no processo de treinamento em comparação aos dados pessoais.

Quando dados pessoais são utilizados para o treinamento de IA, os controladores devem garantir que apenas os dados necessários para as atividades específicas de treinamento sejam tratados. Isso envolve limitar a coleta e o tamanho dos conjuntos de dados contendo dados pessoais ao mínimo necessário para os objetivos pretendidos. Além disso, os controladores devem assegurar que os dados pessoais utilizados sejam precisos, adequados, relevantes e proporcionais aos objetivos do treinamento. A extensão desses requisitos dependerá do contexto específico do processo de treinamento de IA.

Por vezes, o tratamento de dados pessoais não é essencial e o treinamento de um determinado sistema pode ser realizado com dados anonimizados, sintéticos ou desidentificados. Além disso, técnicas como a incorporação de "ruído" podem ser aplicadas para minimizar o possível impacto do tratamento de dados pessoais. Na prática, antes de tratar dados pessoais, os controladores devem avaliar quais as melhores estratégias, técnicas e se dados desidentificados ou anônimos podem alcançar os mesmos resultados. Como alternativa, os controladores podem explorar o uso de dados e conjuntos de dados sintéticos como substitutos para dados do mundo real. Dados sintéticos são gerados artificialmente para replicar as propriedades estatísticas dos dados reais, mas não incluem nenhuma informação identificável sobre indivíduos.

Embora medidas de mitigação emergentes, como dados sintéticos, sejam promissoras para reduzir a dependência de dados pessoais, limitar indevidamente o acesso a dados corre o risco de criar impactos negativos no desenvolvimento de modelos e dificultar esforços para prevenir e mitigar vieses não intencionais[4]. A minimização de dados não significa que apenas pequenos volumes de dados podem ser usados no treinamento de modelos. Em vez disso, a minimização de dados neste contexto pode ser interpretada para exigir um equilíbrio apropriado que reduza a quantidade de dados pessoais usados ao que é necessário ao longo do ciclo de vida de um sistema de IA, permitindo o desenvolvimento de um modelo de alta qualidade e experiência do usuário. Em outras palavras, "minimização de dados" não pode significar usar menos dados do que seria necessário e apropriado para garantir a alta qualidade do modelo de IA específico.

Como afirma a autoridade de proteção de dados do Reino Unido ("ICO")[5], "o princípio da minimização [semelhante ao princípio da necessidade] não significa 'não tratar dado

[4] https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_ai_report_-_hard_issues_and_practical_solutions_01.17.2020.pdf

[5] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-should-we-assess-security-and-data-minimisation-in-ai/?search=minimisation

pessoal'. O segredo é tratar somente os dados pessoais necessários para o seu propósito".

Ainda de acordo com o ICO, o princípio da necessidade não significa que o sistema de IA não poderá tratar dados pessoais. A bem da verdade, esse princípio exige que o agente de tratamento estabeleça quais dados são adequados e relevantes, à luz do uso do sistema de IA no caso concreto.

Os modelos de aprendizado de máquina são projetados para codificar padrões gerais, em vez de fatos específicos sobre exemplos individuais de treinamento, mesmo quando esses exemplos incluem dados pessoais. Essa abordagem permite que os sistemas de IA sejam treinados em conformidade com os requisitos legais, como os princípios e as disposições da LGPD. Ao mesmo tempo, ela garante que a qualidade e a diversidade dos dados usados no treinamento permaneçam essenciais, equilibrando a conformidade legal com a eficácia do sistema de IA.

A partir dessas considerações, verifica-se que é importante que a ANPD reconheça que operações com "quantidades massivas" de dados pessoais não correspondem obrigatoriamente a uma violação ao princípio da necessidade. É essencial que ocorra, portanto, uma análise caso a caso do propósito do tratamento, com base, inclusive, em uma diferenciação entre as aplicações práticas da IA (como, por exemplo, no treinamento do modelo fundacional e no *fine-tuning* posterior do sistema para demandas especializadas), à luz do contexto avaliado.

2.Quais boas práticas e salvaguardas devem ser observadas visando à definição de finalidades específicas e à divulgação de informações claras e adequadas e facilmente acessíveis aos titulares a respeito do tratamento de dados pessoais realizado durante o desenvolvimento e o uso de sistemas de IA?

RESPOSTA: As empresas que tratam dados pessoais para desenvolver sistemas de IA podem identificar uma série de salvaguardas. Por exemplo, se uma empresa conduz uma avaliação de impacto de proteção de dados para atividades relacionadas à IA, a avaliação pode ajudar a identificar potenciais salvaguardas. Somente após a conclusão de uma avaliação de risco, os controladores podem projetar eficazmente sistemas de IA que possibilitem o exercício dos direitos dos titulares de dados e forneçam informações significativas aos titulares sobre os riscos associados ao tratamento de dados pessoais. Esses passos são essenciais para garantir a transparência e lidar de maneira eficaz com os pedidos de acesso dos titulares de dados. Nesse sentido, frameworks como o do NIST ou ISO sobre governança de sistemas de IA são instrumentos de referência bastante relevantes e boas práticas já adotadas pelo mercado, independentemente de obrigações legais ou regulatórias.

O monitoramento contínuo é crucial para sistemas de IA estendendo-se além da análise de risco inicial e configuração de governança de dados. Sistemas de IA, particularmente os de autoaprendizagem, podem desenvolver vieses. Auditorias periódicas simplificam as solicitações de acesso do titular dos dados, oferecendo análises atualizadas de movimentação de dados e ações do sistema. Isso capacita os controladores de dados a lidarem preventivamente com novos riscos, permitindo que eles melhorem e ajustem suas avaliações de risco e planos de governança de dados.

Controladores de dados que usam IA para tratar dados pessoais devem implementar avaliações de risco completas e governança de dados fortes. Isso ajuda a identificar riscos potenciais e determinar se entradas ou saídas de IA constituem dados pessoais. Esse entendimento permite processos de desenvolvimento e sistema de IA que gerenciam efetivamente solicitações de acesso de titulares de dados, permitindo modificações ou exclusões sem interromper a função de IA ou descartar modelos. Isso pode envolver minimizar o uso de dados pessoais, excluir dados confidenciais ou prevenir violações de dados. Saídas consideradas arriscadas ou derivadas de dados confidenciais podem ser removidas proativamente. Auditorias regulares garantem a eficácia contínua das avaliações de risco e estratégias de mitigação.

Essa abordagem flexível permite implementação personalizada com base em circunstâncias individuais, acomodando a natureza diversa e evolutiva da IA. Os objetivos principais, alinhados com os princípios de proteção de dados são entender completamente as implicações de proteção de dados de um sistema de IA específico, comunicar esses riscos aos titulares dos dados de forma clara e transparente e garantir o tratamento eficiente de solicitações subsequentes.

3.Como compatibilizar os princípios da finalidade e da transparência com o uso de sistemas de IA de propósito geral, isto é, sistemas que possam realizar uma ampla variedade de tarefas distintas e servir a diferentes finalidades?

RESPOSTA: Uma organização pode criar um conjunto de dados para treinar um modelo de classificação de imagens (pessoas, veículos, alimentos etc.) e torná-lo publicamente acessível, sem que nenhum uso operacional específico seja previsto durante o desenvolvimento do modelo.

Este modelo pode ser livremente reutilizável, sujeito à sua licença (permitindo adaptações como aprendizagem de transferência) e aos direitos de imagem relevantes e regulamentações de propriedade intelectual. Terceiros podem utilizá-lo para desenvolver sistemas de visão computacional com diversas aplicações, como sistemas de câmera aprimorados para monitoramento de presença na plataforma ou detecção de defeitos em imagens de controle de qualidade do produto.

A finalidade do tratamento, como a realização do tratamento para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades, durante o desenvolvimento se alinha com o princípio da limitação da finalidade quando especificar: (i) o tipo de sistema que está sendo desenvolvido (por exemplo, grandes modelos de linguagem, sistema de visão computacional, IA generativa para imagens/vídeo/áudio), explicado de forma clara e compreensível aos titulares dos dados, apesar da complexidade técnica e dos rápidos avanços; e (ii) funcionalidades e capacidades tecnicamente viáveis, o que significa que o controlador pode elaborar uma lista de capacidades que ele pode prever razoavelmente no estágio de desenvolvimento do sistema.
Dadas as limitações factuais de visibilidade do desenvolvedor sobre a finalidade para a qual os dados pessoais são utilizados em seus sistemas de IA de propósito geral, parece-nos que a responsabilidade pelo princípio da finalidade deve recair sobre a empresa que utilizada dados pessoais nesses sistemas. Assim, no contexto da implementação e uso da IA, a empresa que utiliza o sistema de IA, caso realize tratamento de dados pessoais, deverá ser considerado controlador dos eventuais dados compartilhados por ela, enquanto o desenvolvedor operador.

Caso um modelo venha a ser treinado para fins de ajuste fino, a empresa será controladora se os dados pessoais utilizados no processo forem de sua responsabilidade, e não do desenvolvedor.

Por vezes, o desenvolvedor do sistema de IA de propósito geral pode até mesmo indicar aos contratantes do sistema as finalidades para as quais a implantação deve se destinar e garantir a conformidade. Essa conformidade dependerá, em particular, da observação dessas finalidades pelas empresas contratantes e eventual atualização do seu ROPA, caso haja o tratamento de dados pessoais.

Vale destacar que, de acordo com o *Article 29 Data Protection Working Party* ("WP")[6], dados pessoais podem ser tratados para mais de uma finalidade, sem que isso represente uma violação à legislação de proteção de dados.

Por fim, no que se refere a atividades posteriormente adotadas a partir dos sistemas de IA de propósito geral é necessário que o agente de tratamento atualize seu RoPA, a fim de providenciar informações transparentes e constantemente atualizadas aos titulares a respeito do tratamento de seus dados pessoais. A nível interno, ademais, é essencial que os agentes conduzam avaliações periódicas para identificar se os sistemas estão adequados às finalidades inicialmente delimitadas e descritas no RoPA.

Em suma os desenvolvedores devem ser obrigados a prestar contas de seu papel na construção e no treinamento de modelos de IA e de usos razoavelmente previsíveis, e os

[6] https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf

aplicadores devem ser responsáveis por avaliar os riscos associados às suas implantações específicas dos sistemas e pelo treinamento do sistema, caso de fato o façam para tornar o sistema mais eficaz para o seu uso específico.

4. Quais boas práticas e salvaguardas, bem como parâmetros ou critérios, devem ser considerados ao longo de todo o ciclo de vida de sistemas de IA para prevenir discriminações ilícitas ou abusivas?

RESPOSTA: Os sistemas de IA estão possibilitando novas experiências e habilidades para pessoas ao redor do mundo. Além de recomendar aplicativos, vídeos curtos e programas de TV, os sistemas de IA podem ser usados para tarefas mais críticas, como prever a presença e a gravidade de uma condição médica, combinar pessoas com empregos e parceiros ou identificar se uma pessoa está atravessando a rua. Esses sistemas computadorizados de assistência ou tomada de decisão têm o potencial de ser mais justos e inclusivos em uma escala mais ampla do que os processos históricos de tomada de decisão baseados em julgamentos humanos. O risco é que qualquer viés em tais sistemas também possa ter um impacto em larga escala. Assim, à medida que o impacto da IA aumenta em todos os setores e sociedades, é fundamental trabalhar em direção a sistemas que sejam livres de vieses inclusivos para todos.

Mitigar vieses em modelos de machine learning é desafiador porque esses modelos aprendem com dados do mundo real. Esses vieses podem ser aprendidos inadvertidamente e até mesmo amplificados pelo modelo.

Além disso, garantir equidade em todas as situações e culturas é difícil, mesmo com treinamento e testes extensivos. Um sistema de reconhecimento de fala treinado em adultos brasileiros, embora potencialmente inclusivo nesse contexto, pode falhar em reconhecer a linguagem em evolução dos adolescentes, ter dificuldades com sotaques regionais portugueses ou ter um desempenho ruim com padrões de fala específicos, como gagueira, mesmo dentro de seu público-alvo. O uso pós-lançamento pode revelar resultados imprevistos e injustos que eram difíceis de prever durante o desenvolvimento.

De qualquer forma, é essencial que, para evitar discriminações ilegais e abusivas, os sistemas de IA sejam treinados com dados de qualidade e diversos. Conjuntos de dados insuficientemente diversos levam a saídas tendenciosas de modelos de IA com um impacto discriminatório em indivíduos ou grupos de indivíduos. Os modelos de IA se beneficiam de serem treinados em uma ampla gama de dados para serem úteis para implantação em uma ampla gama de contextos, por exemplo, nos diversos campos de educação e pesquisa. A exposição a um conjunto de dados amplo e diverso também beneficia a sociedade ao reduzir o risco de saídas imprecisas, tendenciosas ou mesmo prejudiciais. É fundamental que os controladores tenham salvaguardas demonstráveis para proteger todos esses direitos fundamentais e mitigar riscos (por exemplo,

realizando avaliações de risco e RIPDs, garantindo a qualidade dos dados, reparação e fornecendo transparência apropriada).

Nesse sentido, é importante que haja flexibilidade na aplicação dos princípios de proteção de dados na fase de desenvolvimento da IA, justamente para permitir um treinamento eficaz que combata vieses e outros cenários adicionais onde os impactos sobre os titulares dos dados podem ser consideravelmente mais significativos.

BLOCO 2 – HIPÓTESES LEGAIS

5.O tratamento de dados pessoais no contexto de sistemas de IA pode ser amparado pela hipótese legal do consentimento? Em quais circunstâncias? Quais as limitações para a utilização dessa hipótese legal nesses contextos e quais salvaguardas devem ser observadas?

RESPOSTA: Existem diversas bases legais previstas na LGPD, e a mais apropriada dependerá do contexto específico. O consentimento pode ser uma base legal apropriada em casos em que os controladores de dados têm um relacionamento direto com as pessoas cujos dados eles querem tratar. No entanto, pode ser difícil coletar consentimento válido para operações de tratamento mais complexas como aquelas envolvidas em IA.

Depender do consentimento também significa permitir que as pessoas revoguem o consentimento, o que é difícil, se não impossível de gerenciar no contexto de IA, e significa que os controladores precisam identificar uma nova base legal se continuarem o tratamento, ou mesmo ter que interromper ou cancelar o funcionamento de determinados modelos de IA, uma vez que os dados pessoais usados em seu treinamento, para os quais o consentimento foi dado, foram revogados e, portanto, quaisquer dados decorrentes de seu tratamento também devem ser excluídos. Em outras palavras, o consentimento não é a base legal mais recomendada para legitimar o tratamento de dados pessoais envolvendo IA, uma vez que sua possível retirada pode comprometer estruturalmente os modelos de IA e impactar significativamente o funcionamento dos sistemas de IA.

Uma eventual obrigação de estabelecer o consentimento como única base legal válida para legitimar o uso de dados pessoais para treinamento de IA poderia inviabilizar a inovação e o aprimoramento dos sistemas de IA, pois tornaria o processo de legitimação mais burocrático. A possibilidade de usar qualquer uma das bases legais é a legalmente mais precisa, além de trazer mais dinamismo para permitir o tratamento legítimo de forma adequada.

Sob a perspectiva do titular dos dados, confiar somente no consentimento também criaria problemas, criando um cenário de fadiga de consentimento, onde o titular dos dados se veria repetidamente notificado e tendo que consentir ou não com o uso de seus dados pessoais para tal propósito, o que é incompatível com a realidade dinâmica da atual economia digital. Reforce-se, inclusive, que outras bases legais, como o legítimo interesse, ao exigir o teste de balanceamento são bases muito mais protetivas dos direitos dos titulares, garantindo que um sopesamento dos direitos e expectativas seja feito pelo agente de tratamento.

Em suma, considerando que a própria LGPD traz um conjunto de bases legais, não nos parece apropriado que uma norma infralegal restrinja o uso de quaisquer bases legais.

6.O tratamento de dados pessoais, no contexto de sistemas de IA, pode ser amparado pela hipótese legal do legítimo interesse? Em quais circunstâncias? Em caso afirmativo, quais salvaguardas devem ser adotadas nessas situações com vistas à proteção de direitos dos titulares, especialmente considerando a vedação de tratamento de dados pessoais sensíveis com base na hipótese legal do legítimo interesse? Em particular, a coleta de dados pessoais para o treinamento de sistemas de IA, especialmente mediante técnicas de raspagem de dados, pode ser fundamentada na hipótese legal do legítimo interesse?

RESPOSTA: O legítimo interesse vem sendo considerado, por autoridades de proteção de dados em diferentes jurisdições, como uma base legal adequada para fundamentar o tratamento de dados pessoais no contexto de sistemas de IA, especialmente para viabilizar o seu treinamento. Com efeito, a autoridade de proteção de dados do Reino Unido (ICO)[7] sustenta que o legítimo interesse corresponde a uma base legal adequada para fins de treinamento de LLM, desde que o agente de tratamento observe o princípio da necessidade.

O legítimo interesse é uma das bases legais previstas no art. 7º da LGPD para legitimar o tratamento de dados pessoais. A confiança em interesses legítimos está, no entanto, sujeita a três condições:

- O interesse perseguido pelo controlador deve ser "legítimo";
- O tratamento deve cumprir a condição de "necessidade";
- O tratamento não deve afetar desproporcionalmente os direitos e interesses dos titulares dos dados, levando em consideração suas expectativas razoáveis. Portanto, é necessário "equilibrar" os direitos e interesses em jogo à luz das condições específicas para sua implementação.

[7] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-do-we-ensure-lawfulness-in-ai/#legitimate-interests

O controlador é obrigado a examinar a conformidade do seu tratamento com essas três condições. Para esse fim, o controlador e o operador devem manter um registro das operações de tratamento de dados pessoais que realizam, especialmente quando baseadas em legítimo interesse (art. 37). Em qualquer caso, onde uma RIPD é necessária, as salvaguardas fornecidas para limitar os possíveis impactos sobre os direitos dos indivíduos devem ser descritas pelo controlador. A realização do teste de balanceamento é um processo fundamental neste sentido, que traz benefícios em termos de visibilidade sobre esses pontos e traz maior clareza sobre os aspectos relevantes do legítimo interesse.

Vale destacar que o § 4º do art. 7 da LGPD prevê a isenção da exigência de consentimento para dados tornados manifestamente públicos pelo titular dos dados. Nesse sentido, é possível afirmar que a coleta de dados pessoais para treinamento de sistemas de IA, inclusive por meio de técnicas de *data scrapping*, pode ser baseada em legítimo interesse. Tanto a Comissão Nacional de Informação e Liberdades (CNIL)[8] francesa quanto o Conselho Europeu de Proteção de Dados (EDPB)[9] sugeriram que o *web scrapping* com base em interesses legítimos pode ser possível.

Além disso, o § 3º do art. 7 da LGPD estabelece que o tratamento de dados pessoais cujo acesso é público deve considerar a finalidade, a boa-fé e o interesse público que justificaram sua disponibilização. A técnica de *web scrapping* pode atender esses critérios, tendo em vista os potenciais benefícios socioeconômicos dos sistemas de IA."

A Brasscom defende, portanto, que dados pessoais, caso estejam publicamente disponíveis ou tenham sido tornados manifestamente públicos pelo titular dos dados, podem ser utilizados, sob o amparo do legítimo interesse, desde que sejam garantidos os direitos do titular e os princípios previstos na LGPD.

Ademais, vale notar que o Guia da ANPD sobre Legítimo Interesse enfatiza a existência de interesses da coletividade que são capazes de embasar operações com dados pessoais. Alinhado a isso, o treinamento de sistemas de IA deve ser analisado também sob a perspectiva da sua relevância social – exemplo disso é o desenvolvimento de modelos voltados à prevenção à fraude.

Considerando as demais bases legais previstas na LGPD, negar a incidência do legítimo interesse significaria, em última instância, inviabilizar o treinamento dos sistemas de IA no contexto brasileiro, posicionando o Brasil na retaguarda do desenvolvimento de tais tecnologias com inúmeras aplicações benéficas.

---

[8] https://www.cnil.fr/en/legal-basis-legitimate-interests-focus-sheet-measures-implement-case-data-collection-web-scraping

[9] https://www.edpb.europa.eu/system/files/2024-05/edpb_20240523_report_chatgpt_taskforce_en.pdf

Nesse sentido, respeitosamente sugerimos que a ANPD reconheça que a coleta de informações publicamente disponíveis para desenvolvimento de IA está dentro das expectativas razoáveis dos indivíduos. A coleta de informações publicamente disponíveis tem sido uma prática comum há muito tempo; ela sustenta grande parte da internet (por exemplo, ela sustenta a indexação usada por mecanismos de busca); e os serviços de IA são amplamente conhecidos e populares. Há uma extensa reportagem sobre IA na grande mídia, e a conscientização e compreensão pública de como esses serviços usam dados publicamente disponíveis para treinar modelos de IA são generalizadas. Além disso, os titulares de dados geralmente entendem que qualquer pessoa pode acessar informações na internet pública, e essas informações não são privadas. Como tal, os titulares de dados não têm as mesmas expectativas de privacidade com relação às informações que escolheram disponibilizar publicamente na Internet.

BLOCO 3 – DIREITOS DOS TITULARES

7. De que maneira os direitos do titular, previstos na LGPD, se aplicam a sistemas de IA?

RESPOSTA: Os direitos dos titulares de dados, bem como a LGPD como um todo, aplicam-se às atividades de tratamento de dados pessoais, independentemente dos meios onde são realizadas. Nesse sentido, a LGPD não se aplica aos sistemas de IA em si, mas sim ao tratamento de dados pessoais realizado por eles.

É importante que a ANPD evite criar requisitos excessivamente específicos sobre os métodos pelos quais um titular de dados exerce esses direitos, devido à grande variedade de controladores que devem honrar solicitações de direitos em uma ampla gama de produtos e serviços. Os melhores métodos de comunicação para exercer esses direitos variam muito entre diferentes produtos e serviços, incluindo aqueles relacionados à IA.

É importante ressaltar que a ANPD deve reconhecer que muitas organizações desenvolverão e usarão sistemas de IA sem tratar dados pessoais — e, portanto, os indivíduos podem não ter, neste contexto, direitos para acessar, corrigir ou excluir dados relacionados a esses sistemas de IA. Por exemplo, se os pesquisadores tratarem dados históricos sobre padrões climáticos para desenvolver e usar um sistema de IA para prever padrões climáticos futuros, eles podem fazê-lo sem processar dados pessoais. Como resultado, os indivíduos não teriam direitos para acessar, corrigir ou excluir dados pessoais processados em conexão com o sistema de IA — uma vez que sua criação e uso não envolveram dados pessoais em primeiro lugar. De forma mais ampla, será importante que os indivíduos entendam o escopo dos direitos fornecidos pela LGPD,

para reduzir a confusão sobre como esses direitos se aplicam quando os sistemas não contêm dados pessoais.

8. Quais as boas práticas e as salvaguardas a serem observadas na disponibilização de canais de atendimento ao titular para exercício dos seus direitos, a exemplo dos direitos de acesso, de oposição e de revisão de decisões automatizadas, no contexto do tratamento de dados pessoais por sistemas de IA? Se possível, descreva as ferramentas utilizadas para implementação de tais canais de atendimento, com os respectivos parâmetros utilizados.

A disponibilização de canais de atendimento ao titular deve observar, da mesma maneira que nos demais cenários de tratamento de dados pessoais, o acesso facilitado por parte dos titulares às informações sobre o tratamento de seus dados, bem como aos contatos de suporte dos agentes, a fim de permitir o exercício de todos os direitos previstos na LGPD. A divulgação dos canais em materiais do agente de tratamento, websites, termos de uso, políticas ou avisos de privacidade continuam a corresponder, nesse cenário, a boas práticas que atendem tanto ao princípio da transparência (artigo 6º, VI, LGPD) quanto à obrigação dos agentes de tratamento em garantir e efetivar os direitos de titulares previstos em lei.

É preciso, contudo, que os direitos de titulares sejam avaliados à luz das limitações técnicas envolvido em determinado tratamento, sendo certo que nenhum direito é absoluto. Nos cenários de impossibilidade da efetivação de determinado direito, o agente de tratamento deve ser facultado a, conforme previsto na LGPD (artigo 18, § 4º, LGPD), indicar as razões de fato e de direito que impedem o atendimento imediato do direito em questão ao titular.

Além disso, é importante e desejável que os canais de atendimento deem ao titular o máximo de autonomia e automação e, na medida do possível, viabilizem a autodeterminação informativa. No entanto, tais canais não podem revelar-se óbices ou embaraços para o exercício de direitos. A adoção de supervisão ou revisão humana é essencial para garantir a efetividade do canal, notadamente nos casos excepcionais ou menos recorrentes.

Os canais de atendimento devem também adotar medidas para garantir a legitimidade do requerente, a fim de não se tornarem instrumentos de usos maliciosos ou não autorizados, nem pontos de vulnerabilidade para incidentes de dados pessoais.

9. Deve haver salvaguardas e limites específicos para o tratamento de dados pessoais sensíveis e para o tratamento de dados pessoais de crianças, adolescentes e idosos durante as etapas do ciclo de vida de sistemas de IA?

RESPOSTA: A Brasscom inicialmente ressalta que a LGPD prevê critérios diferenciados apenas para o tratamento de dados pessoais de crianças e adolescentes, em linha com as premissas trazidas pela ONU, mas a legislação não traz o mesmo entendimento para o caso de dados pessoais de idosos. A norma apenas prevê que cabe à ANPD assegurar que o tratamento de dados dessa parcela da população seja realizado de forma simples, clara, acessível e adequada ao seu entendimento (art. 55-J, inciso XIX, da LGPD). Assim, qualquer tentativa de equiparar os idosos a crianças e adolescentes, ou de classificar seus dados pessoais como dados sensíveis, ultrapassa os limites estabelecidos pela legislação.

No que tange a menores, a LGPD, e os documentos e pareceres emitidos pela ANPD já preveem diversas regras e critérios para o tratamento de dados pessoais de crianças e adolescentes, bem como dados pessoais sensíveis.

Dito isso, vale destacar que, caso o tratamento de dados pessoais de crianças e adolescentes seja realizado com base em legítimo interesse, o controlador terá um patamar maior a ser considerado no processo de elaboração do teste de balanceamento para determinar as melhores medidas de mitigação de risco a serem adotadas.

O teste de balanceamento é efetivamente fundamentado na gestão de risco e *accountability*. Ela permite o tratamento de dados pessoais quando não resulta em risco aos interesses e direitos e liberdades fundamentais dos indivíduos. Ela também promove a proteção dos indivíduos, pois exige que as organizações realizem as avaliações de risco necessárias, definam as medidas de mitigação, treinem os funcionários sobre riscos e medidas de mitigação, monitorem a eficácia contínua das mitigações, identifiquem potenciais lacunas de conformidade, consertem-nas e continuem melhorando o nível de proteção.

Portanto, a flexibilidade fornecida pela base legal de legítimo interesse, juntamente com a responsabilidade organizacional e sua abordagem inerente baseada em risco, torna a base de interesses legítimos um facilitador essencial da IA responsável, inclusive no que tange a dados de menores.

10. Quais os requisitos a serem observados para a garantia e a aplicação do direito à revisão de decisões automatizadas (art. 20 da LGPD)? O que pode ser considerado como decisão tomada unicamente com base em tratamento automatizado de dados pessoais? Quais interesses poderiam ser afetados?

RESPOSTA: As decisões automatizadas oferecem benefícios significativos para a sociedade e para os próprios titulares de dados pessoais, promovendo eficiência, rapidez

e coerência em diversos setores. Como destacado pelo ICO[10], esses mecanismos têm o potencial de transformar áreas como saúde, educação, serviços financeiros e marketing, permitindo soluções inovadoras e acessíveis para desafios complexos.

Sobre os requisitos para a aplicação do direito à revisão de decisões automatizadas, é importante destacar os aspectos estruturantes deste conceito:

(i) deve tratar de decisões automatizadas no contexto da LGPD, ou seja, que envolvem o tratamento de dados pessoais;
(ii) conforme o artigo 20 da LGPD, essas decisões devem ser tomadas unicamente com base em tratamento automatizado, sem intervenção humana; e
(iii) devem afetar os interesses dos titulares de dados pessoais.

Quanto ao segundo questionamento, "decisão tomada unicamente com base em tratamento automatizado de dados pessoais" significa um processo de tomada de decisão que é totalmente automatizado e exclui qualquer influência humana no resultado. Um processo ainda pode ser considerado totalmente automatizado se um humano insere os dados a serem tratados e, então, a tomada de decisão é realizada por um sistema automatizado. Um processo não deve ser considerado totalmente automatizado se alguém pondera e interpreta o resultado de uma decisão automatizada antes de aplicá-la ao indivíduo.

Vale destacar aqui que a lei se refere apenas a decisões, sendo distinta de mera inferência, recomendação e geração de conteúdo, hipóteses nas quais o art. 20 não é aplicável.

Em relação ao terceiro questionamento, para determinar quando um interesse é efetivamente afetado, recomenda-se que a ANPD, considerando o contexto brasileiro, adote critérios ou situações que envolvam:

(i) o impedimento de o indivíduo exercer regularmente seus direitos previstos no ordenamento jurídico, de maneira prolongada, resultando em efeitos negativos que comprometam o exercício de direitos fundamentais;
(ii) o prejuízo no acesso do indivíduo a serviços essenciais, como educação, saúde ou transporte; e
(iii) situações ilícitas que gerem impactos graves ou prejudiciais, como casos de discriminação.

Por fim, a Brasscom reforça a importância de a ANPD reconhecer que tais mecanismos oferecem vantagens significativas para a sociedade como um todo e para os próprios titulares de dados pessoais, não devendo a priori tratar as decisões automatizadas como

[10] https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/individual-rights/individual-rights/rights-related-to-automated-decision-making-including-profiling/

algo necessariamente negativo. Afinal, as decisões automatizadas são essenciais para o desenvolvimento econômico, pois otimizam a conexão entre oferta e demanda, garantindo maior eficiência operacional. A automação facilita a escalabilidade dos serviços digitais, por exemplo, permitindo que as plataformas operem em múltiplos mercados simultaneamente, atendendo a um número crescente de usuários e estabelecimentos sem comprometer a qualidade do serviço.

Além disso, a automação e a inteligência artificial são peças-chave na criação de um ambiente mais seguro e confiável para usuários. Sistemas automatizados podem identificar comportamentos de risco, prevenir fraudes e promover a aplicação consistente de políticas, assegurando que as interações dentro da plataforma sejam justas e transparentes. Esses sistemas também permitem ajustes contínuos com base em dados em tempo real, possibilitando respostas rápidas a variações na demanda ou incidentes específicos. Tais mecanismos tornam o trabalho das plataformas mais acessível e contribuem para fortalecer a economia, ao fomentar um mercado de trabalho dinâmico, inovador e adaptado às necessidades da era digital.

11. Em que hipóteses e sob quais condições pode ser necessária a revisão humana de decisões automatizadas com vistas à adequada garantia de direitos dos titulares?

RESPOSTA: O art. 20 da LGPD estabelece que o titular dos dados tem o direito de solicitar a revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem seus interesses. A LGPD não aventa a possibilidade, ou a necessidade, dessa revisão ser realizada por um humano.

Vale mencionar que o direito à revisão humana de decisões automatizadas - previsto na legislação europeia, mas não recepcionada pela LGPD - não se destina a decisões, ainda que automatizadas, que:

- Executem uma tarefa procedimental específica;
- Sirvam à detecção de padrões de tomada de decisão ou desvios de padrões anteriores de tomada de decisão, sem a intenção de substituir ou influenciar uma avaliação humana previamente concluída;
- Sejam utilizadas em tecnologias tais como:
    1) ○ Prevenção à fraude e à segurança do titular;
    2) ○ Segurança da informação, como anti malware, antivírus, firewall, filtros contra spam e robocalls
    3) ○ Calculadoras, planilhas e bancos de dados;
    4) ○ Registro de domínios, carregamento de sites, gestão de redes, web caching, hospedagem de sites ou qualquer tecnologia similar;
    5) ○ Armazenamento de dados;

6) ○ Matching ou pairing;
7) ○ Correção ortográfica;
8) ○ Comunicação com consumidores em linguagem natural com o objetivo de fornecer informações, fazer indicações ou recomendações, e responder a perguntas (chatbot), desde que sujeita a uma política de uso aceitável que proíba a geração de conteúdo discriminatório ou prejudicial.

12. Quais os parâmetros a serem observados para o fornecimento de informações claras e adequadas a respeito dos critérios e dos procedimentos utilizados para a decisão automatizada, nos termos do § 1º do art. 20 da LGPD? Quais limites e parâmetros de segredo comercial e industrial justificam a não observância do fornecimento de informações, conforme disposto no mesmo dispositivo legal?

RESPOSTA: Sistemas de IA que tomam decisões com base em dados pessoais devem garantir ao titular o direito de entender os critérios e os procedimentos subjacente dessas decisões (art. 20 da LGPD). Isso inclui a obrigação de fornecer informações claras e acessíveis sobre como os dados são tratados. De nossa parte, entendemos que a positivação do direito à explicação na LGPD demanda do controlador que seja sim capaz de fornecer informações compreensíveis para o titular, por meio das quais o titular possa solicitar revisão da decisão automatizada referida no caput do artigo 20 e exercer outros de seus direitos. Além disso, o direito à explicação permite ao titular entender se a decisão automatizada que afeta os seus interesses ofende ou não o princípio da não discriminação. Essa interpretação é suportada pela exigência de que o tratamento de dados deve ser legítimo, explícito, informado e transparente para o titular.

Os segredos comercial e industrial consistem, em si, em parâmetros a serem observados no fornecimento de informações aos titulares, em conformidade com o disposto na LGPD (art. 6º, VI, e art. 55-J, II). Com efeito, o dever de transparência e de fornecer ao titular informações claras e adequadas deve ser compatibilizado com a proteção de informações confidenciais que podem expor empresas à concorrência desleal e à espionagem industrial – em especial no contexto emergente de desenvolvimento de sistemas de IA.

Nesse sentido, as informações a serem prestadas pelos agentes de tratamento devem encontrar limites como (i) as metodologias empregadas em determinado tratamento; (ii) dados específicos e confidenciais relativos ao negócio de determinada empresa, que implicam ou geram valor comercial; (iii) informações de caráter estratégico ou de segurança, que sejam relevantes para o modelo de negócios de determinada empresa ou setor e os quais o/a diferenciam no mercado.

BLOCO 4 – BOAS PRÁTICAS E GOVERNANÇA

13. De que forma programas de governança em privacidade podem ser utilizados como um mecanismo de promoção da conformidade do desenvolvimento e uso de sistemas de IA com a LGPD? Quais requisitos, especificamente relacionados ao desenvolvimento e uso de sistemas de IA, devem ser observados nesses casos?

RESPOSTA: A autoridade de proteção de dados do Reino Unido ("ICO") indica que as implicações da IA para a proteção de dados dependem particularmente dos usos específicos dos sistemas. Para o ICO, contudo, é importante incorporar a proteção de dados desde a concepção e por padrão na cultura e nas atividades de tratamento do agente de IA, processos alinhados com o conceito de *accountability*.

Em termos práticos, tendo em mente o conceito de *accountability*, as organizações devem adotar medidas que implementem requisitos legais de privacidade aplicáveis e que elas sejam capazes de demonstrar a existência e eficácia de tais medidas tanto interna quanto externamente mediante solicitação. Efetivamente, isso significa que as organizações devem implementar programas abrangentes de privacidade e segurança de dados que cubram todos os aspectos do tratamento de dados, incluindo coleta, uso, transferência para terceiros e descarte. Vale ressaltar que medidas de governança de proteção de dados são adotadas independentemente de quais tecnologias são usadas para tratar dados pessoais.

Ter um programa de privacidade abrangente em vigor permite a conformidade com as obrigações legais aplicáveis. No entanto, *accountability* não é apenas sobre conformidade. O programa de privacidade de uma organização pode frequentemente ir além do que é exigido por lei. Para ser eficaz, um programa de privacidade deve ser operado de forma a ter o apoio ativo de todos os funcionários e ser totalmente legitimado dentro da organização pela liderança mais sênior e incorporado à cultura e aos valores éticos da organização. *Accountability* organizacional fornece benefícios significativos a todas as partes interessadas — as próprias organizações, indivíduos e reguladores. Ela permite a conformidade com os requisitos legais e confere confiança e vantagens competitivas às empresas, fornece proteção consistente e eficaz para indivíduos, além de aumentar a transparência organizacional.

14. Considerando o princípio da responsabilização e prestação de contas, quais informações devem ser documentadas durante o ciclo de vida de um sistema de IA? Em quais contextos específicos relacionados a sistemas de IA é recomendada a elaboração de RIPD? Neste caso, é possível estabelecer requisitos específicos a serem observados na elaboração do RIPD?

RESPOSTA: Considerando que a LGPD é um marco legal tecnologicamente neutro, entendemos que não caberia à ANPD estabelecer requisitos específicos para a

elaboração de RIPD envolvendo sistemas de IA. A ANPD pode até incentivar a elaboração de relatórios de impacto quando certas atividades de tratamento de dados pessoais possam acarretar riscos relevantes para os titulares dos dados, conforme avaliação de risco à proteção de dados do controlador, mas não determinar critérios adicionais para sistemas de IA. A LGPD se aplica às atividades de tratamento de dados pessoais, independentemente do meio pelo qual são realizadas, sendo inapropriado criar regras específicas para determinadas tecnologias.

No contexto da privacidade, as informações relacionadas à proteção de dados devem ser documentadas durante o ciclo de vida de um sistema de IA pelas organizações, de acordo com o contexto e sua posição na cadeia de valor, para ajudar a analisar as operações de tratamento planejadas em detalhes e determinar os riscos específicos de privacidade envolvidos e desenvolver estratégias para mitigá-los. As informações podem incluir os tipos de dados coletados, as respectivas finalidades, bases legais, resultados de avaliações de risco, medidas de salvaguarda adotadas, entre outros. Os controladores podem ser encorajados a desenvolver documentação, observando as melhores práticas e padrões voluntários internacionais, como NIST's AI Risk Management Framework[11], e ISO 42001[12] e 23894[13] , o que traria ganhos em termos de alinhamento e interoperabilidade global, fator determinante para a inclusão do Brasil na cadeia global de valor afeta a inteligência artificial.

15. Considerando o ciclo de vida de um sistema de IA, em que momento e contexto do tratamento seria viável ou necessária a anonimização? Qual a técnica utilizada? Quais outras medidas de segurança poderiam ser eventualmente utilizadas visando à proteção da privacidade de titulares de dados?

RESPOSTA: Como afirmamos acima, para mitigar preocupações com privacidade, há muitas medidas de segurança que podem ser adotadas, como o uso de dados anonimizados, sintéticos ou desidentificados em vez de dados pessoais, quando estes não são necessários para cumprir a finalidade identificada, bem como a incorporação de "ruído" para minimizar os dados pessoais que estão sendo tratados. Em outras palavras, antes de coletar dados pessoais, os controladores podem considerar se dados desidentificados ou anônimos podem ser usados para atingir os mesmos resultados. Alternativamente, os controladores também podem considerar se dados sintéticos e conjuntos de dados podem ser usados em vez de coletar dados relacionados a pessoas reais. Dados sintéticos replicam os componentes estatísticos de dados do mundo real, no entanto, eles não contêm identificadores e são gerados artificialmente.

[11] https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

[12] https://www.iso.org/standard/81230.html

[13] https://www.iso.org/standard/77304.html

A Brasscom gostaria de destacar que o artigo 12 da LGPD estipula que os dados anonimizados deixam de ser considerados dados pessoais. Isso implica que, ao seguir os critérios definidos pela legislação para anonimização, os dados perdem sua capacidade de identificar indivíduos, oferecendo uma alternativa viável à eliminação completa dos dados.

Portanto, a LGPD reconhece tanto a anonimização quanto à eliminação como medidas equivalentes para garantir a proteção dos dados pessoais, oferecendo às organizações a flexibilidade necessária para escolher a abordagem mais adequada às suas necessidades e à natureza dos dados envolvidos.

Algumas outras medidas como criptografia, controle de acesso, uso de pseudonimização, auditorias regulares, limitação de acesso por contexto, e ambientes seguros para processamento e revogação de acesso de tempos em tempos também deve ser considerada como boa prática e instrumentos que viabilizam a proteção e privacidade dos titulares de dados.

Devido à complexidade dos processos de treinamento de sistemas de IA, a Brasscom entende que seria interessante que a ANPD adotasse o entendimento de que esforços de anonimização de dados, por desenvolvedores nos processos de treinamento, não sejam considerados atividades de tratamento de dados pessoais e, consequentemente, reconhecer que o desenvolvimento do sistema não envolveu dados pessoais. Isso porque o processo de anonimização trará uma camada maior de proteção de dados e garantirá que o processo de treinamento do sistema de IA não trará consequências aos titulares dos dados. Essa flexibilidade incentivará ainda mais a adoção de medidas de segurança para garantir maior privacidade dos titulares dos dados.