Contribuição do IP.rec para a “Tomada de subsídios: Inteligência Artificial e revisão de decisões automatizadas” **da ANPD**

Instituto de Pesquisa em Direito & Tecnologia do Recife

FICHA TÉCNICA

REALIZAÇÃO

Instituto de Pesquisa em Direito e Tecnologia do Recife - **IP.rec**

PESQUISA E TEXTO

Helton Leyendecker

Luana Batista

Rhaiana Valois

COORDENAÇÃO

André Fernandes

REVISÃO

Clarissa Mendes

André Fernandes

PROJETO GRÁFICO

Estúdio PUYA!

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?

Acomodar satisfatoriamente o treinamento de sistemas de IA que demandam grande quantidade de dados no bojo normativo da LGPD em sua forma corrente é impraticável. O princípio da necessidade impõe que o tratamento seja realizado com o mínimo de dados suficientes à realização de sua finalidade. Dentre seus objetivos: evitar excesso de dados tratados; manter a proporcionalidade entre a coleta e a finalidade; e reduzir impactos negativos ao titular, minimizando danos.

Apesar de efetivo sob a ótica da coleta e tratamento individual, o princípio da necessidade esboroa no contexto de treinamento de IA de "larga escala", nos quais se pressupõe quantidades massivas de dados para treinamento. Nestes casos, se a implementação de uma IA for legitimada para uma finalidade válida, orientada por uma base legal, notadamente o legítimo interesse, é certo que haverá necessidade de grandes volumes de dados para operacionalizar esse sistema.

Por essa razão, a aplicação alargada do princípio da necessidade (art. 6º, III, LGPD) como um dos fundamentos a orientar e autorizar o funcionamento destes sistemas de IA tende à permissividade e inconsistência normativa, esvaziando sua função protetiva.

Uma interpretação restritiva do princípio orientaria que o atendimento da "necessidade" se estabelece entre o cotejo da polaridade "titular de dados" e da polaridade "aplicação no tratamento", mas há uma defesa de que se atenderia ao

princípio na medida em que “a aplicação do sistema” “exige” “grandes quantidades de dados” e, portanto, seria “necessário” para o tratamento conforme.

Nota-se que fica questionada a necessidade de utilização de dados nestes sistemas para alcançar a finalidade pretendida, na forma de minimização do uso de dados, quando a finalidade puder ser alcançada de outras formas. O tema, inclusive, tem sido alvo de debates entre especialistas: a busca de resultados similares com modelos que usam menos dados[1].

A título de salvaguardas, registre-se: forma de coleta de dados e sua previsão legal; limitação do uso à finalidade específica; filtragem de dados para garantir qualidade e diversidade; implementação de tecnologias de aprimoração de privacidade (PETs/PPTs); algoritmos de minimização de dados para reduzir o dataset; transparência e accountability organizacional, com práticas contínuas de avaliação e auditoria, permitindo o monitoramento de conformidade.

A implementação eficaz de PETs e PPTs, por outro lado, além dos obstáculos técnicos e de escalabilidade, não é absolutamente eficaz frente às técnicas de reidentificação[2]. O controle de qualidade e diversidade dos dados requer mecanismos de monitoramento robustos. A limitação do uso à finalidade específica se torna desafiadora quando confrontada com inferências ou geração de dados novos[3].

---

[1] DORRIER, Jason. These Mini AI Models Match OpenAI With 1,000 Times Less Data. SingularityHub, 2024. Disponível em: https://singularityhub.com/2024/10/04/these-mini-ai-models-match-openai-with-1000-times-less-data/. Acesso em: 20 jan. 2025.

[2] TSCHIDER, Charlotte. AI's Legitimate Interest: Towards a Public Benefit Privacy Model. *Houston Journal of Health Law & Policy*, v. 21, p. 153, 2021. Disponível em: https://ssrn.com/abstract=3725933. Acesso em: 20 jan. 2025.

[3] SOLOVE, Daniel J. Artificial Intelligence and Privacy. Florida Law Review, v. 77 (a publicar em jan. 2025). GWU Legal Studies Research Paper No. 2024-36, GWU Law School Public Law Research Paper No. 2024-36. p. 33-34. Disponível em: https://ssrn.com/abstract=4713111. Acesso em: 20 jan. 2025.

Em última instância, é possível adotar medidas mitigadoras, mas a LGPD apresenta limitações substanciais para compatibilizar as complexidades do treinamento de IA de larga escala ao princípio da necessidade.

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?

No que se refere à finalidade, a LGPD estabelece que o tratamento de dados pessoais deve ser realizado a partir de propósitos legítimos, específicos e ser informado ao titular, não sendo possível a realização de tratamento posterior incompatível com as finalidades informadas previamente (art. 6º, I). Objetivando atender a esse princípio, é fundamental que o tratamento de dados pessoais observe, em todas as fases do ciclo de funcionamento da IA, a finalidade específica informada.

Nesse sentido, são pontos de atenção: a definição de finalidade específica que observe as hipóteses legais, não podendo ser ampla, vaga ou genérica, bem como seu informe ao titular; a limitação do tratamento à finalidade definida, mitigando o uso secundário de dados pessoais, que ocorre quando os dados são tratados para propósitos diferentes do inicial; a análise, em casos excepcionais, da compatibilidade, que pode ou não existir entre a finalidade inicial e a finalidade para uso secundário[4]; a observância ao princípio da necessidade, evitando o tratamento de dados desnecessários à consecução da finalidade.

[4] WIMMER, Miriam. Limites e possibilidade para o uso secundário de dados pessoais no poder público: lições da pandemia. Revista Brasileira de Políticas Públicas, v. 11, n. 1, p. 125. 2021. Disponível em: https://www.gti.uniceub.br/RBPP/article/view/7136/pdf#. Acesso em: 19 ago. 2024.

Ademais, o princípio da finalidade deve ser aplicado tanto aos dados coletados quanto às inferências geradas a partir deles[5], assegurando que as informações derivadas sejam tratadas de maneira compatível com os propósitos inicialmente definidos, de modo a evitar usos desproporcionais ou desvios não informados que possam comprometer os direitos dos titulares.

Acerca da responsabilidade do controlador, esta não se limita à definição da finalidade inicial. Deve-se garantir que qualquer alteração nas finalidades do tratamento seja igualmente informada ao titular, com a devida justificativa e reenquadramento nas hipóteses legais. Para tanto, mecanismos de controle e auditoria devem ser implementados para assegurar que o tratamento de dados esteja sempre em conformidade com as finalidades informadas.

Nesse contexto, a transparência em relação à finalidade é fundamental. O titular deve ser informado sobre as finalidades do tratamento, de forma que entenda como seus dados serão utilizados. Em casos de uso secundário ou reutilização de dados, a compatibilidade entre a finalidade inicial e a nova finalidade deve ser rigorosamente verificada, levando em consideração não apenas os dados coletados, mas também as implicações que o uso secundário pode ter para os direitos do titular.

Os critérios estabelecidos pelo Grupo de Trabalho do Artigo 29, atual EDPB, podem ser úteis para verificar se a reutilização de dados é compatível com a finalidade original. Com base nisso, é importante verificar se há conexão entre a finalidade inicial e as novas pretendidas; analisar o contexto de coleta, bem como a natureza dos dados em questão, com atenção especial para informações consideradas sensíveis nos termos da lei; e mensurar as possíveis consequências

[5] WACHTER, Sandra; MITTELSTADT, Brent. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. Columbia Business Law Review, n. 2, p. 22-28. 2019. Disponível em: https://ssrn.com/abstract=3248829. Acesso em: 20 jan. 2025.

do novo tratamento para os titulares de dados, adotando as medidas de segurança adequadas[6].

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?

É preciso considerar que a inclusão de dados de uma pessoa específica em um conjunto de treinamento pode gerar inferências negativas sobre o grupo de pessoas do qual esse indivíduo faz parte, além do uso indevido dessas informações, inclusive por terceiros. Nesses casos, é importante identificar se a utilização dessas informações será feita para identificar padrões e correlações gerais com finalidades meramente estatísticas ou se serão usadas para a criação de perfis individuais[7]. Essa distinção importa, pois os dados usados para fins estatísticos, tornam-se, em geral, dados agregados.

No primeiro caso, a adoção de medidas de segurança robustas pode mitigar os riscos associados ao tratamento de dados pessoais, desde que não utilizados para finalidades incompatíveis com os interesses do grupo em questão. Já na segunda hipótese, é preciso considerar que as inferências feitas sobre os titulares particulares, bem como a eventual reidentificação destes, podem causar danos direitos a pessoas específicas[8]. Em relação a isso, o art. 20 da LGPD é particularmente interessante por possibilitar revisão de decisões automatizadas, permitindo que a Autoridade Nacional realize auditorias para verificação de aspectos discriminatórios nesse tratamento.

---

[6] Grupo de trabalho do artigo 29.º (2018). Orientações relativas à transparência na aceção do Regulamento 2016/679 (WP260 rev.01). Disponível em: https://www.edpb.europa.eu/system/files/2023-09/wp260rev01_pt.pdf. Acesso em: 20 jan. 2025.

[7] The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. European Parliament. 25 jun. 2020. Disponível em: https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641530.

[8] The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. European Parliament. 25 jun. 2020. Disponível em: https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)641530.

Além disso, a LGPD tem um amplo arcabouço principiológico que deve ser observado durante o processamento dessas informações. Nesse sentido, é importante reputar não só as inferências sobre os titulares, mas também a possibilidade de reidentificação a partir de dados agregados, como criação de novos dados pessoais, recaindo sobre estes as mesmas salvaguardas garantidas pela lei para os dados de entrada.

Sobre este último, o art. 12, §2º, prevê que dados anonimizados, quando o processo ao qual foram submetidos puder ser revertido, poderão ser considerados como dados pessoais dos titulares, se estes puderem ser identificados. É preciso observar, no entanto, que o perfilamento pode violar os interesses dos titulares, sem que seja necessário identificar exatamente seu titular. Por isso, adotar uma interpretação mais ampla desse dispositivo é necessária para proteger de maneira mais efetiva os titulares.

É preciso ainda considerar que a reutilização de dados para finalidades distintas daquelas originalmente informadas ao titular pode ser incompatível com a LGPD. Nesse cenário, como mencionado acima, os critérios estabelecidos pelo Grupo de Trabalho do Artigo 29 podem ser úteis para verificar se a reutilização de dados é compatível com a finalidade original[9].

Por fim, em relação à transparência, são necessárias ponderações sobre o funcionamento desses sistemas, já que suas operações internas não são tão simples de serem compreendidas. Nesse sentido, além do inciso VI do art. 6º da LGPD, os arts. 20, § 1º, e o art. 38 da lei contém previsões importantes para embasar o direito à explicação das decisões por sistemas de IA e exigir a produção de relatórios de transparência para os atores envolvidos na cadeia.

[9] Grupo de trabalho do artigo 29.º (2018). Orientações relativas à transparência na aceção do Regulamento 2016/679 (WP260 rev.01). Disponível em: https://www.edpb.europa.eu/system/files/2023-09/wp260rev01_pt.pdf. Acesso em: 20 jan. 2025.

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?

De acordo com o guia "Guidance on AI and Data Protection" do Information Commissioner's Office (ICO)[10], a conformidade exige atenção a aspectos como justiça, transparência e mitigação de vieses, princípios que, no ordenamento nacional, complementam a proteção de dados pessoais prevista na LGPD[11], especialmente nos fundamentos de finalidade, necessidade e não discriminação.

No início do ciclo de vida, a formulação do objetivo e design do sistema são cruciais para garantir que as escolhas algorítmicas não perpetuem desigualdades. As decisões tomadas nessa etapa, incluindo a definição de critérios de desempenho, devem ser acompanhadas de uma avaliação dos riscos éticos e jurídicos associados. Além disso, desde a fase de concepção, o sistema deve ser projetado para minimizar o uso de dados pessoais, garantindo que apenas informações estritamente necessárias sejam processadas. Em sistemas que envolvem decisões automatizadas com impacto significativo sobre os titulares, é recomendável a realização de RIPDs para identificar potenciais riscos e propor medidas mitigadoras.

Na etapa de coleta e tratamento de dados, é indispensável que os dados coletados sejam representativos da diversidade da população impactada. Dados desbalanceados podem introduzir desigualdades sistêmicas nos resultados gerados pelos modelos. Assim, devem ser aplicadas estratégias de análise de dados que identifiquem potenciais vieses, desbalanceamentos ou lacunas que possam resultar em tratamento desigual de grupos específicos.

---

[10] Information Commissioner's Office (ICO). Guidance on AI and Data Protection. Disponível em: https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/. Acesso em: 20 de jan. 2025.

[11] Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm Acesso em: 18 de jan. 2025.

Durante o treinamento e desenvolvimento do modelo, é importante realizar avaliações contínuas para identificar potenciais vieses ilícitos no desempenho do modelo. A utilização de auditorias regulares e avaliações externas é essencial para avaliar o comportamento dos sistemas. Além disso, deve-se adotar estratégias que identifiquem vieses não apenas nas variáveis diretamente relacionadas a características sensíveis, mas também em variáveis correlacionadas. Essas práticas ajudam a detectar discriminações que, de outra forma, poderiam ser mascaradas pela complexidade do modelo.

Na fase de operação, o monitoramento contínuo é uma salvaguarda essencial para detectar e corrigir eventuais comportamentos discriminatórios emergentes. Nessa fase, também são recomendados processos de auditoria regulares e a manutenção de registros detalhados (documentação técnica); isso é fundamental para garantir a rastreabilidade e a transparência, bem como para assegurar a responsabilização em caso de impactos adversos. Ademais, a revisão humana em decisões automatizadas que afetem significativamente os direitos dos titulares deve ser garantida, conforme previsto na LGPD.

Por fim, quando um sistema de IA atinge o fim de sua vida útil ou é substituído, é essencial garantir que os dados tratados sejam protegidos de usos indevidos através de um processo verificável e auditável.

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?

No contexto de sistemas de IA, o tratamento de dados pessoais com base no consentimento enfrenta limitações significativas, devido à complexidade e opacidade inerentes a esses modelos. Embora o consentimento seja reconhecido entre técnicos e leigos como uma das principais bases legais para o tratamento de dados, sua aplicabilidade no âmbito de sistemas de IA é questionável.

Em relação ao ato de consentir, é razoável pressupor que a informação deva ser compreendida, avaliada e ponderada antes que se emita um veredicto acerca do consenso ou dissenso. No entanto, a dificuldade em termos de explicabilidade das operações realizadas por sistemas de IA torna o ato de consentir altamente oneroso para o titular[12], que é confrontado com a tarefa de compreender como seus dados serão utilizados nesse sistema e o que isso pode acarretar. Aceitar que o consentimento devidamente ponderado é a exceção implica aceitar que o manejo desta base legal estaria, como regra, comprometido por vício.

Além da explicabilidade, é latente o problema das inferências. Mesmo quando o titular consente com o uso de seus dados, o tratamento pode gerar inferências que extrapolam o escopo do consentimento inicial, expondo os titulares a situações e riscos não consentidos. Dessa forma, consentir com o tratamento de dados para uso em IA equivale, em muitos casos, a conceder uma autorização geral que pode ser explorada de maneiras que os titulares não conseguem antecipar ou compreender. É o que a professora Elettra Bietti chama de "passe livre" em sua obra "Consent as a Free Pass". [13]

O problema das inferências evidencia a necessidade de assegurar que estas também estejam sujeitas a salvaguardas legais adequadas[14], tendo em vista a inexistência de mecanismos legais para endereçá-las, contestá-las ou retificá-las. Essa lacuna jurídica reforça a fragilidade do consentimento enquanto salvaguarda, uma vez que os titulares não têm controle significativo sobre como seus dados serão reinterpretados nos sistemas de IA.

---

[12] CORREN, Ella. The Consent Burden in Consumer and Digital Markets. *Harvard Journal of Law & Technology*, v. 36, n. 2, p. 568-576, Spring 2023. Disponível em: https://jolt.law.harvard.edu/assets/articlePDFs/v36/Corren-Consent-Burden.pdf. Acesso em: 20 jan. 2025.
[13] BIETTI, Elettra. Consent as a Free Pass: Platform Power and the Limits of the Informational Turn. *Pace L. Rev.*, v. 40, p. 308, 385. 2020. Disponível em: <https://ssrn.com/abstract=3489577>. Acesso em: 20 de jan.de 2025.
[14] WACHTER; MITTELSTADT, *op. cit.*, p. 81.

Ademais, a impossibilidade prática de exercer direitos como a revogação de consentimento no contexto de treinamento de IAs amplia as limitações dessa hipótese[15]. Dessa forma, o consentimento, além de ser difícil de implementar de forma verificável e rastreável em sistemas de IA, torna-se insuficiente para proteger os titulares em um ambiente onde decisões automatizadas podem restringir escolhas ou impor resultados sem a participação ativa dos indivíduos.

Portanto, o consentimento como base legal, em sua forma atual, é insuficiente para proteger os titulares no contexto de sistemas de IA. Mais do que aprimorar sua rastreabilidade, o foco deve estar na construção, adaptação e aprimoramento do arcabouço regulatório existente, de modo a limitar práticas abusivas e promover a transparência e a explicabilidade no tratamento de dados.

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?

Para que o legítimo interesse possa ser utilizado como base legal no contexto de sistemas de IA, é necessário que sejam preenchidos certos requisitos e que sejam adotadas medidas mais rigorosas para proteger os direitos dos titulares de dados. De acordo com art. 10 da LGPD, essa base legal somente pode ser aplicada quando o controlador possuir propósitos legítimos, considerados a partir de

[15] YOO, Christopher S.; WOLFF, Josephine; LEHR, William. Lessons from GDPR for AI Policymaking. Virginia Journal of Law & Technology, v. 27, p. 22, 2024. Disponível em: https://scholarship.law.upenn.edu/cgi/viewcontent.cgi?article=1376&context=faculty_articles. Acesso em: 20 jan. 2025.

situações concretas, sem que isso cause um impacto significativo aos direitos dos titulares e que suas expectativas sejam consideradas como parte da avaliação de interesses legítimos (Legitimate Interest Assessment - LIA).

Nessas circunstâncias, é preciso que o controlador demonstre que o tratamento de dados é essencial para alcançar os fins legítimos pretendidos, priorizando os interesses e as legítimas expectativas dos titulares de dados, que são a parte mais vulnerável nesse processo. Além de realizar o registro de todas as etapas do processamento de dados, é fundamental que o tratamento de dados para o treinamento de IA observe os princípios da LGPD, bem como minimize, na medida do possível, a coleta de dados para esse fim. Diante disso, é importante verificar se o uso do dado tem o potencial de violar os direitos dos titulares a partir de avaliações periódicas de risco.

Quando a coleta de dados envolver raspagem de dados, é importante ainda que haja garantias de que isso foi realizado de forma ética e transparente e que dados sensíveis ou que se refiram a crianças e adolescentes não sejam incorporados durante a coleta. Nesse ponto, convém destacar que essa base legal não deve ser usada para fundamentar o tratamento desses tipos de dados, tendo em vista o risco que o uso inadequado dessas informações pode acarretar para os seus titulares. É preciso ainda adotar precauções adicionais quando a raspagem envolver conteúdos protegidos por direitos autorais e conexos, garantindo que o titular desses direitos seja informado e receba a justa remuneração pelo uso desse material.

Nesse sentido, a atuação da Autoridade é fundamental para fiscalizar e monitorar o uso dessas informações pelas empresas e pelo poder público. Um exemplo disso foi a decisão tomada no ano de 2024, quando a Autoridade suspendeu preventivamente a vigência da nova Política de Privacidade da Meta. Essa suspensão foi fundamentada na ausência de base legal apropriada para

justificar a coleta de dados pessoais para treinamento do seu modelo de IA generativa[16].

É preciso considerar ainda que essas tecnologias permitem a reutilização desses dados para propósitos distintos daquele para o qual os dados foram inicialmente coletados, podendo ser utilizados, dessa forma, para finalidades mais específicas. Nesses casos, é preciso distinguir cada operação de processamento e identificar seus propósitos específicos, de modo que seja adotada a base legal que melhor se adequa ao caso em questão, tendo em vista à proteção dos direitos dos titulares de dados.

[16] BRASIL. Autoridade Nacional de Proteção de Dados. Voto nº 11/2024/DIR-MW/CD. Brasília, DF: Autoridade Nacional de Proteção de Dados, 02 jul. 2024.

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

Os direitos previstos na LGPD aplicam-se aos sistemas de IA, regulando o tratamento de dados pessoais em todo o ciclo de vida e mitigando riscos à privacidade, discriminação e segurança.

Neste contexto, a LGPD estabelece que os titulares sejam informados de maneira clara e acessível sobre os tratamentos de seus dados, incluindo os processos subjacentes. Esta transparência requer a implementação de práticas que garantam a explicabilidade, permitindo que os titulares compreendam as regras de negócio e os critérios utilizados para gerar previsões ou tomar decisões automatizadas. Isso inclui a criação de *logs* auditáveis, a adoção de modelos interpretáveis e a documentação detalhada de algoritmos e parâmetros. Para modelos mais complexos, como LLMs, é recomendável uma documentação abrangente que descreva os dados de treinamento, os objetivos e as limitações do sistema.

Além disso, para garantir a conformidade, é necessário assegurar a revisão humana das decisões críticas geradas pelos modelos, especialmente em casos de decisões automatizadas com impactos significativos, como crédito e saúde[17]. A LGPD assegura ao titular o direito fundamental de questionar os critérios adotados e solicitar uma revisão humana. Para atender a essa exigência, devem ser implementados processos robustos que viabilizem intervenções humanas e garantam que as informações sobre as bases de decisão sejam compreensíveis.

[17] General Data Protection Regulation (GDPR). Disponível em: https://gdpr-info.eu/. Acesso em: 13 de jan. de 2025.

Ademais, os direitos de retificação, exclusão e anonimização possuem implicações diretas na governança de sistemas de IA, assegurando que os titulares possam corrigir, eliminar ou anonimizar seus dados. Assim, os sistemas devem incorporar mecanismos que permitam ajustes dos dados e evitem a retenção de informações desnecessárias, podendo demandar a reconfiguração do modelo ou o reprocessamento de dados.

No que tange ao consentimento, ressalvadas as problemáticas expostas antes, é necessário implementar sistemas que documentem o consentimento de maneira verificável, com capacidade de rastrear as finalidades específicas para as quais os dados foram coletados. Particularmente em contextos de grandes volumes de dados, é crucial manter políticas que respeitem as limitações impostas pelo titular, inclusive em casos de revogação.

Para garantir a efetividade do direito à portabilidade, são necessárias soluções técnicas que assegurem a interoperabilidade entre sistemas, permitindo que os titulares transfiram suas informações de forma segura e eficiente. Esta necessidade se traduz na adoção de padrões abertos e APIs que viabilizem o compartilhamento de dados sem comprometer a integridade ou confidencialidade.

Por fim, os sistemas de IA devem incorporar, desde a fase de projeto, a realização de RIPDs, especialmente em casos de tratamentos de alto risco[18]. Essas avaliações contribuem para a identificação e mitigação de riscos técnicos, como ataques adversariais, vazamento de dados e a perpetuação de vieses, além da implementação de controles de segurança rígidos.

[18] EU Artificial Inteligence Act. The AI Act explorer. Disponível em: https://artificialintelligenceact.eu/ai-act-explorer/. Acesso em: 15 de jan. de 2025.

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.

Primeiramente, a possibilidade do reconhecimento de inferências como dado pessoal deve ser considerada, ainda que, inicialmente, através de interpretação extensiva do art. 12, §2º da LGPD. Isso é essencial para garantir que os titulares tenham controle não apenas sobre os dados fornecidos, mas também sobre as informações geradas a partir desses dados, como perfis ou predições. Daniel J. Solove, em *Artificial Intelligence and Privacy* (2024), assevera que enquanto a maioria das leis de privacidade dá o direito aos indivíduos de retificarem seus dados ou consentirem com sua coleta, elas normalmente são falhas em permitir que os titulares retifiquem ou se oponham a inferências derivadas de seus dados[19].

Dessa forma, os canais de atendimento devem incluir mecanismos que possibilitem ao titular obter informações detalhadas sobre as inferências realizadas, bem como os critérios, dados e modelos utilizados para gerá-las. Tais canais devem permitir que o titular questione a validade e a razoabilidade dessas inferências, solicitando sua revisão, especialmente em situações onde decisões automatizadas impactem significativamente seus direitos. Ademais, devem ser estruturados de forma acessível, tanto em termos de linguagem quanto de usabilidade, com interfaces que sejam inclusivas e responsivas.

Neste ponto, esbarramos nos obstáculos da explicabilidade e da disponibilidade de informações transparentes e coesas, principalmente ao tratar de inferências. É essencial implementar mecanismos que traduzam conceitos

[19] SOLOVE. *op. cit.*, p. 33.

complexos como: recursos visuais, materiais interativos e atendimento humanizado. Esses canais devem não apenas informar, mas também capacitar o titular a compreender o impacto das inferências sobre seus direitos, incentivando-o a corrigir ou se opor a informações que considere imprecisas ou inadequadas.

No tocante à disponibilização das inferências e de suas bases, inclusive lógicas, os principais entraves enfrentados são, sem dúvidas, a alegação de segredo comercial por parte das companhias e a dificuldade técnica de segmentar os dados gerados para cada titular, disponibilizando-os e categorizando-os de forma individualizada. O primeiro entrave, de caráter regulatório, pode ser contornado por meio de *enforcement* jurídico robusto, hermenêutica da da legislação vigente e regulamentações que padronizem tais requisições.

Já o segundo, de natureza técnica, pressupõe a adoção de tecnologias como *dashboards* personalizados, integrados a sistemas de IA Explicável (xAI) para apresentar de forma visual e compreensível os dados e inferências relacionados a cada titular. Além disso, o uso de APIs padronizadas para exportação de informações e de modelos interoperáveis contribuiria para garantir a transparência necessária no processo de tratamento de dados inferidos. No entanto, para que isso seja possível, é necessário que os modelos de IA sejam projetados ou adaptados para permitir esse nível de explicabilidade e rastreabilidade, o que nem sempre é trivial.

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?

É fundamental que o tratamento de dados pessoais sensíveis receba um tratamento mais restrito, bem como haja a limitação das bases legais previstas na lei. Nesse sentido, é preciso considerar que informações sensíveis podem ser extraídas e inferidas de dados pessoais que, à primeira vista, não apresentam uma relação direta com categorias definidas na lei. Isso aumenta o risco de que tais

dados sejam utilizados de maneira discriminatória, muitas vezes sem que isso seja sequer percebido.

Nessas circunstâncias, a utilização do consentimento deve ser reconsiderada, pois, além da dificuldade inerente de explicabilidade desses modelos, é preciso considerar a opacidade intencional que permeia o desenvolvimento da IA, de modo que não há como falar em consentimento informado, muito menos específico no contexto de IAs de propósito geral.

Além disso, tendo em vista as vulnerabilidades sociais e econômicas e assimetrias informacionais existentes, é provável que os titulares não consigam avaliar adequadamente os riscos associados à anuência com o tratamento de dados e cedam seus dados sem maiores considerações. Um exemplo disso é a investigação anunciada pela ANPD contra uma empresa suspeita de oferecer dinheiro em troca do escaneamento da íris dos indivíduos[20].

Dessa forma, é importante que inferências sobre categorias de dados sensíveis e a eventual reidentificação dessas informações seja considerado como tratamento de dados pessoais sensíveis, recaindo sobre esse processamento as mesmas obrigações e medidas protetivas.

Nesse sentido, uma abordagem de risco é essencial para lidar com essa questão, de modo que seja possível estabelecer e antecipar medidas de segurança. Além disso, é fundamental ampliar a possibilidade de exercício coletivo desses direitos, tendo em vista não só a efetividade das medidas adotadas, como também as assimetrias de poder existentes entre os titulares e as empresas.

---

[20] MAGALHÃES, André Lourenti. ANPD investiga práticas de empresa que coleta íris em troca de dinheiro. Disponível em: https://canaltech.com.br/seguranca/anpd-investiga-praticas-de-empresa-que-coleta-iris-em-troca-de-dinheiro/. Acesso em: 24 jan. 2025.

No que tange aos dados de crianças e adolescentes, apesar da LGPD trazer uma seção específica sobre o tratamento desses dados, essas informações não estão incluídas no rol de dados sensíveis previsto no art. 5º, II. Tendo em vista a condição de vulnerabilidade no qual esse público se encontra, é imprescindível tratar essas informações como sensíveis no contexto de uso por sistemas de IA, além de trazer limitações adicionais sobre o uso desses dados.

Nesse sentido, é fundamental observar os princípios do ECA, minimizar a coleta de dados, restringir seu uso para fins como publicidade infantil e adotar medidas robustas de segurança para mitigar os riscos associados ao desenvolvimento de IA nesse contexto.

Por fim, em relação aos dados de pessoas idosas é preciso, de forma similar, considerar a vulnerabilidade desses indivíduos e adotar as salvaguardas necessárias para garantir a sua proteção. Por essa razão, é importante considerar também que informações referentes a esses indivíduos podem ser sensíveis em determinados contextos.

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?

O art. 20 da LGPD prevê aos titulares o direito à revisão de decisões automatizadas, permitindo que solicitem a revisão de decisões que os afetem significativamente e que tenham sido tomadas exclusivamente com base no tratamento automatizado de dados pessoais. Para assegurar a aplicação desse direito, é necessário observar uma série de requisitos. Primeiramente, o controlador deve oferecer informações transparentes e compreensíveis sobre a lógica, os critérios e os parâmetros utilizados nas decisões automatizadas, assegurando a

rastreabilidade e o registro das variáveis e etapas decisórias. Essa transparência deve ser adaptada para que indivíduos sem conhecimento técnico possam compreender os processos envolvidos.

Além disso, quando uma revisão humana for conduzida, ela deve ser realizada por profissionais qualificados, com autoridade para alterar a decisão e que compreendam o contexto decisório, considerando tanto os dados utilizados quanto possíveis vieses algorítmicos. O PL 2338/2023[21] reforça a importância da supervisão humana efetiva como elemento central para garantir a revisão de decisões automatizadas. Nesse sentido, a supervisão deve ser significativa, ou seja, não pode ser meramente simbólica ou protocolar.

É igualmente essencial que os titulares sejam informados, de forma clara, sobre o direito à revisão e que tenham acesso a mecanismos para exercer esse direito, como canais de comunicação simples e eficazes. Por fim, medidas de mitigação de riscos devem ser implementadas para prevenir vieses, erros e discriminações, protegendo os titulares de decisões automatizadas que possam ser injustas ou prejudiciais.

No contexto da GDPR, uma decisão automatizada é caracterizada como aquela baseada exclusivamente no tratamento automatizado de dados pessoais para chegar ao resultado final, sem qualquer intervenção humana. Isso implica que os dados são processados por algoritmos ou sistemas de IA que determinam o resultado ou a recomendação final (na forma de tomada de decisão out the loop).

Os interesses potencialmente afetados por decisões automatizadas abrangem uma ampla gama de aspectos fundamentais, econômicos, sociais e reputacionais. Decisões enviesadas podem comprometer o direito à igualdade e à

[21] Brasil. Senado Federal. Projeto de Lei n. 2338, de 2023. Substitutivo aprovado em 10 de dez. de 2024. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em: 20 de jan. de 2025.

não discriminação, afetando de forma desproporcional grupos vulneráveis, como mulheres ou minorias raciais. Também há riscos econômicos, como a negativa de crédito ou a criação de perfis imprecisos que prejudiquem a empregabilidade ou o acesso a serviços essenciais. Do ponto de vista da privacidade, o uso inadequado de dados sensíveis e a falta de transparência podem minar a confiança dos titulares e expô-los a riscos adicionais. A autonomia individual é igualmente vulnerável, pois decisões automatizadas podem restringir escolhas ou impor resultados sem o consentimento explícito do titular. Além disso, desigualdades preexistentes podem ser agravadas, especialmente no caso de populações com menor acesso ou familiaridade com tecnologias digitais.

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?

A revisão humana de decisões automatizadas é uma condição essencial para garantir a adequada proteção dos direitos dos titulares de dados, especialmente quando essas decisões podem ter consequências significativas e potencialmente prejudiciais. O AI Act e a GDPR, assim como o PL 2338/2023, estabelecem que, em determinados contextos, as decisões automatizadas são passíveis de revisão humana, principalmente quando podem afetar substancialmente os direitos, liberdades e interesses dos indivíduos.

No AI Act, a revisão humana é prevista para sistemas de alto risco, como aqueles que envolvem decisões automatizadas com impacto significativo sobre os direitos dos indivíduos. O titular tem o direito de solicitar essa revisão, que visa mitigar riscos de erros, discriminação ou injustiças que possam surgir dos processos automatizados.

A exigência de revisão humana em sistemas de alto risco está diretamente vinculada aos princípios de transparência e explicabilidade. Segundo o AI Act, a

transparência não se limita a informar os titulares sobre os dados processados: deve fornecer clareza sobre o processo decisório, incluindo a lógica do algoritmo, os dados utilizados e os critérios decisórios. A revisão humana, quando solicitada, deve ser conduzida por profissionais qualificados que tenham acesso a essas informações e capacidade de identificar erros ou vieses nos algoritmos.

Em paralelo, o PL 2338/2023, também enfatiza a supervisão humana de decisões automatizadas em sistemas de alto risco. A legislação proposta enfatiza a necessidade de supervisão humana efetiva para garantir a proteção dos direitos dos titulares, tornando obrigatória a revisão quando solicitada pelo titular.

A revisão humana é fundamental em diversas situações críticas, principalmente quando as decisões automatizadas têm impacto jurídico na vida do titular. É essencial para mitigar riscos de discriminação e preconceito oriundos de vieses nos dados de treinamento dos modelos de IA. Além disso, a revisão é necessária quando há suspeita de erros ou inconsistências nos resultados do algoritmo ou, conforme estabelecido na LGPD, quando o titular solicitar uma revisão.

Ademais, é importante que o processo de revisão seja auditável, permitindo que as decisões revisadas sejam comunicadas de forma clara e transparente ao titular, garantindo sua compreensão sobre os motivos pelos quais a decisão foi mantida ou alterada. A revisão humana de decisões automatizadas deve ser vista como uma ferramenta de mitigação de riscos. Ao permitir que uma pessoa avalie o contexto da decisão automatizada, reduz-se o risco de consequências adversas, como a aplicação indevida de critérios discriminatórios.

A revisão não deve ser apenas um procedimento administrativo, mas uma instância real de correção no processo de desenvolvimento, em que se possa questionar os parâmetros algorítmicos, exigir ajustes nos dados utilizados e até reverter decisões que possam estar em desacordo com os direitos dos titulares.

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?

Em relação ao segredo comercial e industrial, algumas considerações merecem ser destacadas. Em primeiro lugar, não há uma definição legal expressa para o termo na legislação brasileira. A Lei de Propriedade Industrial (LPI) trata do tema no art. 195, inciso XI, como crime de concorrência desleal e utiliza critérios como utilidade, não publicidade e não obviedade da informação ou conhecimento para delimitar sua incidência.

Além disso, como signatário do Acordo sobre Aspectos dos Direitos de Propriedade Intelectual Relacionados ao Comércio (TRIPS) de 1994, o Brasil adota os requisitos previstos no art. 39 desse tratado, que incluem o sigilo da informação, a consideração do seu valor econômico com base na sua manutenção em segredo e a implementação de medidas, pelo titular, para garantir sua confidencialidade. Com base nisso, a doutrina tenta delimitar o conceito, mas é preciso evitar que o termo seja utilizado para inviabilizar o exercício de direitos, especialmente no que diz respeito à transparência e à explicabilidade das decisões.

Segundo Ana Frazão, em artigo publicado no JOTA, na LGPD já é possível identificar as bases para limitar o uso indiscriminado dessa proteção por parte das empresas[22]. Isso porque no art. 20 da lei é previsto o direito dos titulares à explicação e à revisão das decisões automatizadas, assim como é estabelecido que, em caso de recusa de divulgação das informações com base na proteção do

[22] FRAZÃO, Ana. Transparência de algoritmos x segredo de empresa. JOTA. 09 jun. 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/transparencia-de-algoritmos-x-segredo-de-empresa. Acesso em: 20 jan. 2025.

segredo comercial e industrial pela empresa, a ANPD é competente para realizar auditorias sobre os aspectos discriminatórios em tratamento automatizado de dados pessoais. Nesse mesmo sentido, a LPI também prevê limitações ao dispor sobre a possibilidade de divulgação confidencial em processos judiciais, desde que sob segredo de justiça (art. 206).

O art. 6º do PL 2338/2023 apresentava uma relevante previsão que obrigava a disponibilização de explicações suficientes, adequadas e inteligíveis a qualquer pessoa ou grupo afetado por um sistema de IA que produzisse efeitos jurídicos relevantes ou de alto risco, ao mesmo tempo que exigia informações, como as características de funcionamento do sistema, os dados processados e suas fontes, os critérios utilizados e os mecanismos disponíveis para contestação.

Ademais, como aspectos discriminatórios são difíceis de prever e até mesmo identificar, é importante estimular a realização de auditorias externas independentes, como uma forma de antecipar eventuais usos discriminatórios dos modelos e garantir maior confiabilidade às informações disponibilizadas. Somado a isso, é preciso considerar, mais uma vez, a dimensão coletiva de proteção dos dados pessoais, por meio de ações coletivas para reivindicar eventuais injustiças provocadas.

Tendo em vistas as assimetrias existentes, é preciso ainda considerar a possibilidade de inversão do ônus da prova, bem como da adoção da presunção relativa em situações complexas com interesses legítimos conflitantes[23].

[23] FRAZÃO, Ana. Transparência de algoritmos x segredo de empresa. JOTA. 09 jun. 2021. Disponível em: https://www.jota.info/opiniao-e-analise/colunas/constituicao-empresa-e-mercado/transparencia-de-algoritmos-x-segredo-de-empresa. Acesso em: 20 jan. 2025.

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?

Os programas de governança em privacidade são instrumentos importantes para promover a conformidade de sistemas de IA com os princípios e normas da LGPD. Esses programas desempenham um papel fundamental na proteção dos dados pessoais e, em última análise, na garantia dos direitos fundamentais dos indivíduos.

Em relação aos requisitos específicos, é essencial que esses programas estabeleçam obrigações sobre a qualidade e integridade dos dados, avaliem a sua relevância e necessidade conforme o contexto de aplicação da IA, bem como assegurem a conformidade dos protocolos de acesso e tratamento de dados. Além disso, é preciso implementar controles de acesso rigorosos, de modo a evitar o vazamento e o uso indevido dessas informações.

O AI Act da União Europeia, por exemplo, propõe a segmentação de sistemas de IA com base no nível de risco que representam, estabelecendo mais obrigações para os modelos considerados mais perigosos. No Brasil, o PL 2338/2023 adota estratégia semelhante. O art. 29 do projeto, por exemplo, exige a realização de uma avaliação preliminar do modelo de propósito geral que deverá considerar as finalidades de usos esperadas, bem como identificar os respectivos níveis de risco esperados e potencial risco sistêmico[24]. Somado a isso, o art. 30

[24] Brasil. Senado Federal. Projeto de Lei n. 2338, de 2023. Substitutivo aprovado em 10 de dez. de 2024. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em: 20 de jan. de 2025.

dispõe de uma série de requisitos adicionais que deverão ser observados antes que o modelo possa ser disponibilizado no mercado[25].

É necessário ainda que o programa disponha sobre medidas específicas de governança de dados, como a filtragem das fontes de dados, realização de avaliações contínuas e de testes e simulações para combater ataques, além da implementação de escudos de *prompts* e da filtragem de conteúdo que possa afetar direitos.

Nesse sentido, é importante salientar que a governança deve abranger tanto os desenvolvedores *upstream* quanto os *downstream*, os quais devem ser informados sobre as obrigações legais e técnicas que precisam observar no desenvolvimento de sistemas mais específicos que utilizem como base um modelo fundacional.

Nesse contexto, é indispensável que os desenvolvedores *upstream* implementem um sistema de gestão de dados para assegurar o cumprimento das obrigações, bem como mantenham os registros sobre processamento dos dados. As obrigações devem ser distribuídas conforme os papéis desempenhados por esses atores e as responsabilidades pelos eventuais danos causados a terceiros devem ser compartilhadas entre eles.

Por fim, o programa de governança deve incluir obrigações de transparência, não apenas em relação aos resultados gerados pela IA, especialmente em sistemas generativos, que precisam ser devidamente rotulados como tal, mas também quanto à origem das informações, às etapas de desenvolvimento, as decisões automatizadas realizadas, além lógica básica envolvida nesses resultados. Para isso, é essencial também a realização de auditorias periódicas e a publicação dos relatórios produzidos nesse sentido.

---

[25] Brasil. Senado Federal. Projeto de Lei n. 2338, de 2023. Substitutivo aprovado em 10 de dez. de 2024. Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233. Acesso em: 20 de jan. de 2025.

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?

O princípio da responsabilização e prestação de contas demanda que informações detalhadas sejam documentadas em todas as etapas do ciclo de vida de um sistema de IA, incluindo planejamento, treinamento, validação, implementação, monitoramento e descontinuação. É essencial que as informações documentadas abranjam desde a forma de coleta dos dados pessoais até o armazenamento, utilização, exclusão e produtos resultantes desses dados, especialmente nos casos em que permitam reconstrução e reidentificação de seus titulares.

A forma de coleta, especialmente em sistemas de IA, deve ser claramente detalhada, incluindo a origem dos dados (se são coletados diretamente dos titulares, por meio de fontes públicas ou via raspagem de dados) e os métodos utilizados, como a automação ou coleta por terceiros. Esses registros devem ser auditáveis, possibilitando a verificação de conformidade com as hipóteses legais de coleta e a comprovação de que as medidas de proteção e privacidade estão sendo seguidas.

Além disso, devem estar documentadas a finalidade do sistema, as bases legais para o tratamento de dados em todas as suas variações, os critérios para seleção e limpeza dos dados, a lógica dos algoritmos de tratamento e inferências, as métricas de desempenho, os resultados dos testes de robustez e explicabilidade, as estratégias de mitigação de riscos e os incidentes relacionados à privacidade ou à segurança.

Acerca do RIPD, no contexto dos sistemas de IA, recomenda-se que sua elaboração seja a regra e não a exceção. Isso se deve ao fato de que os sistemas de IA frequentemente envolvem tratamentos de dados pessoais, podendo representar riscos elevados para as liberdades civis e os direitos fundamentais dos titulares. Conforme estabelece o Art. 5º, XVII, da LGPD, o RIPD deve documentar os processos de tratamento que possam gerar esses riscos, bem como as medidas, salvaguardas e mecanismos de mitigação adotados.

No caso dos sistemas de IA, os riscos incluem discriminação, a tomada de decisões automatizadas, o uso (ou geração) de grandes volumes de dados sensíveis, ou o impacto da coleta de dados por meio de raspagem, que conforme aponta Solove em *Artificial Intelligence and Privacy*, ignora princípios fundamentais da privacidade, de modo que dados são simplesmente coletados por terceiros, sem qualquer aviso, consentimento, triagem, salvaguarda, propósito específico, limitação de finalidade, minimização, respeito aos direitos individuais, etc[26]. Em suma: a prática de raspagem de dados é desprovida de quaisquer considerações de privacidade.

Dessa forma, o RIPD deve abranger, no mínimo, os tipos de dados, a metodologia de coleta, a forma de obtenção e a análise do controlador quanto às medidas de segurança implementadas para proteger esses dados. Finalmente, o RIPD deve detalhar como os dados serão tratados e armazenados ao longo de todo o ciclo de vida do sistema de IA, abordando medidas como a anonimização, criptografia, e controles de acesso.

[26] SOLOVE. *op. cit.*, p. 25.

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?

No ciclo de vida de um sistema de IA, a anonimização de dados é essencial para proteger a privacidade dos titulares e garantir a conformidade com a LGPD. Conforme regulações como o GDPR e o AI Act da União Europeia, a anonimização é particularmente necessária nas etapas de coleta, armazenamento, processamento e compartilhamento de dados, especialmente quando se lida com informações sensíveis.

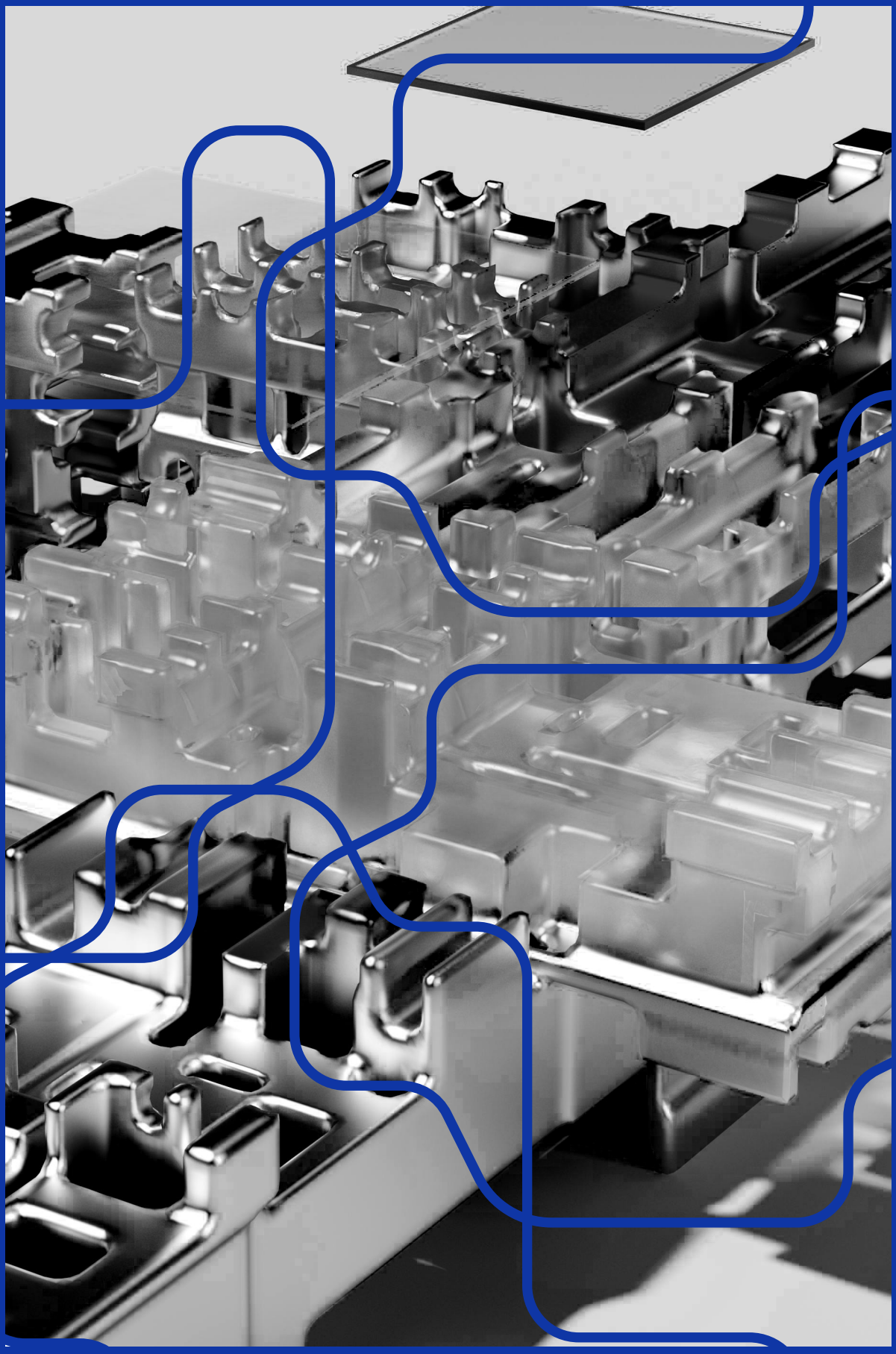
O princípio da minimização de dados previsto no GDPR estabelece que sejam capturadas apenas informações relevantes para os objetivos definidos, assim como a LGPD. Para implementar a anonimização de forma eficaz, diferentes técnicas podem ser aplicadas de acordo com o contexto e os objetivos do tratamento, como a generalização, a randomização, a supressão, a tokenização, a perturbação, além da utilização de dados sintéticos. Cada técnica deve ser escolhida considerando sua eficácia na proteção da privacidade e seu impacto na utilidade dos dados.

A anonimização pode ser aplicada já na fase de coleta para evitar exposição de dados pessoais, reduzindo riscos de identificação durante as fases seguintes. No armazenamento de dados, medidas adicionais podem ser aplicadas para proteger as informações e evitar acessos não autorizados. A descentralização, por meio de métodos como o aprendizado federado, permite que os dados permaneçam nos dispositivos dos usuários, evitando sua centralização em grandes servidores. Técnicas como a pseudoanonimização também desempenham um papel crucial nesse estágio, substituindo informações sensíveis por códigos reutilizáveis que mantêm a utilidade dos dados para análise.

Durante o processamento, o desafio está em equilibrar a utilidade dos dados com a proteção da privacidade. A aplicação de técnicas de anonimização, como a perturbação, ajuda a modificar os dados de forma que se preservem propriedades estatísticas necessárias para análises precisas, enquanto a reidentificação se torna extremamente difícil. A privacidade diferencial surge como uma solução robusta para assegurar que os resultados dos modelos não dependam excessivamente dos dados de qualquer indivíduo específico. Além disso, a seleção criteriosa de variáveis evita o processamento de dados desnecessários.

No compartilhamento e arquivamento de dados, a anonimização torna-se ainda mais crítica para proteger a privacidade dos titulares, garantindo que informações armazenadas para fins estatísticos ou de auditoria não possam ser rastreadas até indivíduos específicos.

Além das medidas citadas, outras medidas de segurança podem ser utilizadas para proteger os dados, como a aplicação dos RIPDs previstos na LGPD, a utilização da criptografia e a adoção de práticas de privacidade e segurança por design. Além disso, devem ser adotadas medidas de monitoramento contínuo para identificar e corrigir vulnerabilidades, bem como treinamento das equipes para prevenir erros e fortalecer a segurança organizacional.



Instituto de
Pesquisa em
Direito & Tecnologia
do Recife