Direito,
Inovação
& Tecnologia

IA Generativa e posicionamentos do EDPS e EDPB

Proteção de Dados Pessoais

Artigo

Rony Vainzof, Jean Santana, Alexandra Krastins e Nuria Baxauli

05.06.2024

www.vlklaw.com.br

IA GENERATIVA E POSICIONAMENTOS DO EDPS E EDPB - PROTEÇÃO DE DADOS PESSOAIS

A governança da Inteligência Artificial (IA) Generativa têm sido objeto de especial atenção de órgãos de proteção de dados na União Europeia. Nas últimas duas semanas, em particular, houve dois posicionamentos relevantes: **(i)** Guia do *European Data Protection Supervisor* (EDPS) sobre a conformidade da IA Generativa; e **(ii)** Relatório do *European Data Protection Board* (EDPB) sobre o ChatGPT.

Seguem os principais pontos de atenção para nortear também empresas brasileiras:

I – GUIA DO EDPS SOBRE O DESENVOLVIMENTO E USO DE SISTEMAS DE IA GENERATIVA

O **Guia do EDPS,** de 03 de junho de 2024, oferece instruções práticas para o desenvolvimento e utilização de sistemas de IA Generativa por instituições europeias. No entanto, o seu conteúdo pode auxiliar à conformidade de qualquer instituição com as normas de privacidade.

Relatório de Impacto à Proteção de Dados

É essencial realizar Avaliações de Impacto de Proteção de Dados (DPIA) antes de operações de alto risco, o que incluí novas tecnologias como a IA Generativa. Agentes de IA, de acordo com suas responsabilidades, devem buscar o aconselhamento do encarregado de proteção de dados (DPO) e tomar medidas para mitigar os riscos identificados. É importante também contemplar a visão dos afetados pelo sistema, e revisar periodicamente os riscos identificados durante todo o ciclo de vida do sistema de IA generativa, documentando todas as decisões e ações tomadas.

Apontamentos sobre o papel do DPO

O DPO deve desempenhar o papel de aconselhamento e assistência independente. Ele deve conhecer como o sistema trata dados pessoais e se envolve em processos de tomada de decisão automatizada. Além disso, deve garantir que o uso do sistema se encontra devidamente documentado e que a transparência adequada foi ofertada ao titular. O EDPS reforça, ainda, o papel do DPO em assistir na elaboração do Relatório de Impacto à Proteção de Dados (RIPD), se necessário, e recomenda, em acordo com as boas-práticas, que seja feito um inventário de uso de sistemas de IA Generativa.

Base legal - GDPR

Embora não limite a aplicação de nenhuma base legal, o EDPS aborda com maior profundidade duas:

- **Legítimo interesse,** o qual o EDPS aponta como base legal potencial para a coleta e uso de dados para fins de treinamento, teste e validação do sistema, desde que respeitado o teste de balanceamento**[1]**.

- **Consentimento,** na qual afirma que, embora seja possível, seu uso deve se operar de forma cuidadosa, sobretudo no contexto de treinamento de IA, considerando, inclusive, o uso de informações publicamente disponíveis. Além disso, se o consentimento for retirado pelo titular, o tratamento anteriormente conduzido com base nele continua legal, desde que tenha sido feito em conformidade com o GDPR. Mas o controlador deve parar de conduzir essas operações de tratamento a partir da sua revogação. Se não houver outra base legal, os dados devem ser deletados. Apesar do EDPS não abordar de forma aprofundada essa questão, esse ponto, sob a perspectiva técnica, é bastante crítico, uma vez que seria necessário retreinar o sistema de IA para que a revogação consentimento seja efetiva.

Minimização de dados

É possível garantir este princípio no âmbito da IA. Os controladores devem garantir que os dados pessoais envolvidos no desenvolvimento de modelos de IA Generativa sejam objeto de controles técnicos para minimizar o uso de dados em todas as fases de desenvolvimento. Devem avaliar, que os dados utilizados são de alta qualidade, sendo apropriadamente rotulados e curados, por meio de processos apropriados de governança, incluindo a revisão periódica e sistemática das bases de dados. Não obstante, as bases de dados e os modelos devem ser acompanhados de documentação apropriada sobre a sua estrutura, manutenção e utilização pretendida. Na contratação com terceiros, também deve ser levado em conta o princípio da minimização.

Acuracidade dos dados

Para garantir a acuracidade dos dados, o EDPS recomenda que as bases de dados de treino tenham sua estrutura e conteúdo verificado por meio de conjuntos de dados para validação durante o treinamento, dentre outras técnicas. Caso esses dados sejam adquiridos de terceiros, cláusulas contratuais apropriadas para garantir sua acurácia devem ser implementadas, assegurando a adoção de medidas que resguardem a qualidade dos dados – como procedimentos apropriados de coleta de dados, procedimentos de preparação, como anotação, rotulagem, limpeza, enriquecimento e agregação, e identificação de lacunas.

Os resultados também devem ser monitorados para garantir a sua qualidade. Ademais, procedimentos de teste e validação do sistema, com as respectivas bases para tanto, devem ser adotados. A checagem recorrente é importante porque sistemas treinados com dados acurados também podem produzir alucinações.

Transparência

Desenvolvedores e utilizadores de IA Generativa devem fornecer informações apropriadas sobre o tratamento de dados pessoais.

Os desenvolvedores devem fornecer informações sobre as diferentes etapas do processo de desenvolvimento, incluindo a origem dos conjuntos de dados, o procedimento de curadoria/rotulagem, bem como qualquer outro associado com o tratamento de dados durante o ciclo de vida da IA.

No uso de sistemas que interajam diretamente com seres humanos, o EDPS recomenda que estes sejam informados que estão interagindo com uma IA.

Tomada de decisões automatizadas

O uso de IA generativa não implica necessariamente a tomada de decisões automatizadas. Isso vai depender do uso e do peso das informações na tomada de decisão final. Quando a IA Generativa for utilizada para tomada de decisões automatizadas, será necessário fornecer ao titular informações sobre a lógica do sistema, o que deve incluir não apenas o funcionamento do algoritmo, mas informações sobre as bases de dados utilizadas.

Ademais, apenas devem ser adotadas tais soluções, se o controlador se encontrar apto a garantir a implementação do direito de revisão humana no processo de tomada de decisão automatizada.

Transferência internacional

Caso o uso do sistema envolva transferência internacional de dados, por exemplo, quando o fornecedor de IA se localize em país terceiro, esse compartilhamento deverá estar em conformidade com uma das bases legais para transferência internacional previstas no GDPR**[2]**.

II – RELATÓRIO DO EDPB SOBRE O CHATGPT

O **Relatório do EDPB,** de 23 de maio de 2024, ressaltou que impossibilidades técnicas não podem ser utilizadas para justificar a não conformidade com o GDPR e apresentou avaliação preliminar em cinco pontos: **(i)** licitude do tratamento; **(ii)** equidade do tratamento; **(iii)** transparência; **(iv)** acuracidade dos dados; **(v)** direitos dos titulares.

Licitude do Tratamento

O EDPB avaliou a legalidade do tratamento de dados em cinco etapas:

1. Coleta de dados de treinamento, incluindo a raspagem de dados da internet.

2. Pré-processamento dos dados de treino, incluindo sua filtragem.

3. Treinamento do modelo.

4. Prompts e outputs do ChatGPT.

5. Treinamento do ChatGPT com *prompts*.

Nas três primeiras etapas, o EDPB destacou os riscos mais relevantes aos direitos e liberdades dos titulares, pois as informações coletadas da internet podem conter informações relevantes sobre a vida privada, incluindo dados pessoais sensíveis.

A OpenAI justificou o tratamento de dados pessoais com base no legítimo interesse – a qual ainda é objeto de investigação pelo EDPB. A despeito disso, o EDPB enfatizou a importância de salvaguardas para minimizar os impactos nos titulares, como evitar a coleta de dados de certas fontes (por exemplo, perfis em redes sociais) e medidas para deletar ou anonimizar dados antes da fase de treinamento, para garantir que o teste de balanceamento penda em favor dos interesses do controlador ao invés do impacto nos direitos fundamentais dos titulares.

Caso dados pessoais sensíveis sejam coletados, o EDPB sugeriu a possibilidade de enquadramento de "dado tornado manifestamente público pelo próprio titular" (art. 9/2(e), do GDPR). No entanto, nem todos os dados publicamente acessíveis foram publicados pelo próprio titular, o que exige práticas adicionais para garantir a conformidade com o GDPR, como a filtragem de categorias de dados durante e após a coleta.

Para as demais hipóteses, a OpenAI também defendeu o uso do legítimo interesse e informou a possibilidade de *opt-out* da utilização de prompts para o treinamento.

Equidade do Tratamento

O EDPB destacou que a OpenAI não pode impor contratualmente aos usuários a obrigação de conformidade com o GDPR. Cabe à OpenAI assegurar sua própria conformidade com o regulamento.

Transparência

O EDPB ressaltou a importância de os titulares serem informados sobre a possibilidade de uso de suas interações com o ChatGPT para treinamento de maneira clara e demonstrável.

Por outro lado, no que se refere à raspagem de dados, devido à impraticabilidade de fornecer essa informação aos titulares, pode-se aplicar a exceção ao dever de informação prevista no art. 14º/5(b) do GDPR**[3]**, já o uso dos dados coletados da interação com o ChatGPT deve ser devidamente informado ao titular.

Direitos dos Titulares

O EDPB reconheceu que a OpenAI já oferece mecanismos que possibilitam o exercício de direitos, inclusive nas configurações das contas dos usuários. No entanto, enfatizou a importância de aprimorar as medidas para atender aos direitos dos titulares,

especialmente considerando as dificuldades técnicas no exercício do direito de correção dos dados, que a OpenAI sugere, em certas situações, substituir pelo direito de eliminação.

Acuracidade dos Dados

A acuracidade dos dados foi dividida em *input* e *output*. Embora a acuracidade seja importante em ambas as etapas, o EDPB reconheceu que, para inputs, o objetivo do tratamento é o treinamento de um sistema probabilístico e não a seleção de informação factualmente precisas. Quanto aos outputs, os usuários tendem a esperar que a informação seja precisa. Por isso, o EDPB destacou que cabe ao controlador informar adequadamente sobre a natureza probabilística dos sistemas e a falta de confiabilidade quanto à acuracidade dos *outputs*, por mais que o princípio da acurácia deva ser privilegiado.

Referências:

1 *Já há precedente favorável para a base legal do legítimo interesse no treinamento de IA, conforme decisão da Autoridade de Dados da Bélgica, a qual, ao avaliar a reclamação de uma titular relativo ao uso de seus dados por uma Instituição Financeira para treinar um Modelo de IA e ofertar descontos personalizados a produtos/serviços de terceiros, tendo entendido por satisfeitos os requisitos do teste de balanceamento. Consulte a decisão: https://www.gegevensbeschermingsautoriteit.be/publications/beslissing-ten-gronde-nr.-46-2024.pdf*

2 *Artigos 44 e seguintes do GDPR.*

3 *Article 14. Information to be provided where personal data have not been obtained from the data subject. 5. Paragraphs 1 to 4 shall not apply where and insofar as: (b) the provision of such information proves impossible or would involve a disproportionate effort, in particular for processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, subject to the conditions and safeguards referred to in Article 89(1) or in so far as the obligation referred to in paragraph 1 of this Article is likely to render impossible or seriously impair the achievement of the objectives of that processing. In such cases the controller shall take appropriate measures to protect the data subject's rights and freedoms and legitimate interests, including making the information publicly available;*

Autores:

Rony Vainzof
VLK Avogados

Jean Santana
VLK Advogados

Alexandra Krastins
VLK Avogados

Nuria Baxauli
VLK Advogados