

Onde o **Direito**
impulsiona a **inovação**

GENERAL-PURPOSE AI CODE OF PRACTICE

First Draft

Como o trabalho da Comissão
Europeia pode servir de modelo
global, inclusive para o Brasil



www.vlklaw.com.br

Introdução

Especialistas independentes apresentaram o primeiro draft do Código de Boas Práticas de IA de Uso Geral (GPAI), de acordo com o EU AI Act. Levaram em conta contribuições dos provedores de IA de uso geral e abordagens regulatórias internacionais, sob os seguintes temas:



Transparência



Direitos autorais



Taxonomia de riscos sistêmicos; e



Medidas técnicas e de governança para mitigação de riscos.

Essa versão inicial será discutida com aproximadamente mil partes interessadas nas próximas semanas, em uma sessão plenária do AI Office da Comissão Europeia. Esse processo de elaboração iterativa começou em 30 de setembro e prevê quatro rodadas de redação até abril de 2025.

Importância do Código

O Código final, previsto no art. 56 do Regulamento, servirá como uma referência **fundamental para o desenvolvimento e implantação responsável de modelos de IA de uso geral confiáveis e seguros**. Além de fornecer diretrizes claras de transparência e segurança, ele proporcionará, na prática, maior segurança jurídica aos fornecedores de GPAI, uma vez que a adesão ao Código servirá como “**presunção de conformidade**” com as suas obrigações, conforme Considerando 117. Essa presunção será especialmente importante durante o período de transição entre a entrada em vigor de tais regras (02/08/2025) e o prazo final para conformidade plena (02/08/2027).

Para provedores de modelos de IA que apresentam riscos sistêmicos, o Código **detalhará uma taxonomia dos riscos e exigirá medidas de avaliação e mitigação**, além de estratégias de governança para atenuar impactos negativos potenciais.



As regras do Código, de acordo com o AI Act, incluem **diretrizes detalhadas para assegurar que os provedores de IA de propósito geral** (incluindo aqueles com risco sistêmico) **operem com segurança, transparência e em conformidade com a legislação da UE**. A seguir, um resumo detalhado das principais regras e medidas listadas no documento.



1) Regras para Provedores de IA de Propósito Geral

A) TRANSPARÊNCIA (ARTIGO 53)



O objetivo é garantir a clareza sobre o uso e as capacidades do modelo de IA para todos os envolvidos - tanto para autoridades quanto para empresas e pessoas que utilizarão o modelo.

- **Documentação para Autoridades (Artigo 53(1)(a)):** Os provedores precisam manter documentos atualizados sobre como o modelo foi treinado e testado, incluindo dados usados, estrutura do modelo (como está organizado internamente), métodos de distribuição (se está disponível na nuvem ou download, por exemplo), e data de lançamento. Isso ajuda a garantir que as autoridades possam avaliar a segurança e conformidade do modelo quando necessário.
- **Documentação para Clientes (Provedores Downstream) (Artigo 53(1)(b)):** Provedores que vão usar o modelo em outras aplicações (como uma empresa que use IA para atendimento ao cliente) precisam receber detalhes sobre as capacidades do modelo e limitações. Essa transparência evita que os clientes usem a IA para algo que ela não foi projetada e ajuda a manter a segurança.
- **Política de Uso Aceitável (AUP):** Um documento claro sobre como o modelo pode ser usado corretamente e o que é proibido, o qual deve incluir:

Escopo: Definição de quem a política abrange e quais recursos ela cobre.

Usos e usuários principais: Definição clara dos usos e usuários pretendidos para o modelo.

Usos inaceitáveis: Detalhamento das ações proibidas com o modelo.

Medidas de segurança: Descrição dos protocolos de segurança que os usuários devem seguir ao utilizar os sistemas de IA de propósito geral.

Monitoramento e privacidade: Explicação sobre por que e como o provedor de IA monitora o uso do modelo e o impacto na privacidade dos usuários.

Processo de advertência e critérios de suspensão: Definição de critérios para suspensão ou retirada de privilégios de usuários que não cumprirem a AUP.

Critérios para encerramento de contas: Referências a leis e regulamentações aplicáveis para reforçar e encerrar contas de usuários, se necessário.

Reconhecimento: Requerimento de que os provedores downstream reconheçam ter lido, entendido e aceitado cumprir a AUP.

B) DIREITOS AUTORAIS (ARTIGO 53(1)(C))

- **Política de Direitos Autorais e Conformidade com a Lei:**



Política Clara e Específica: Os provedores devem estabelecer uma política formal que descreva como os dados utilizados para o treinamento e desenvolvimento de modelos de IA respeitam os direitos autorais. A política precisa incluir uma análise dos dados para garantir que qualquer conteúdo protegido por direitos autorais seja utilizado de maneira legal e que os dados estejam livres de infrações.

- **Respeito ao Protocolo de Exclusão de Robôs (robots.txt):**

Respeito à Configuração de Sites e Restrições: O Código prevê que os sistemas de coleta de dados para treinamento de IA devem respeitar as configurações de robots.txt dos sites, uma prática que informa os provedores sobre as permissões ou restrições de uso de dados.

Rejeição a Fontes de Dados Não Confiáveis: Para evitar violações de direitos autorais, os provedores devem evitar o uso de dados de sites não autorizados ou de fontes piratas, garantindo assim que apenas dados autorizados sejam incluídos no treinamento da IA. Esta medida previne o uso de conteúdo protegido sem a devida autorização e promove o uso ético de dados.

- **Transparência de Direitos Autorais:**

Listagem e Divulgação de Fontes de Dados: Para facilitar a auditoria e assegurar a conformidade com as leis de direitos autorais, o Código exige que os provedores mantenham uma lista detalhada das fontes de dados utilizadas no treinamento de modelos de IA. Isso inclui referências a bases de dados, sites ou outros repositórios de conteúdo, com especificação de qualquer licença ou autorização para uso dos dados.

Detalhamento do Licenciamento: Além de listar as fontes, os provedores devem fornecer detalhes sobre as licenças e permissões obtidas para utilizar conteúdos protegidos por direitos autorais. Isso é essencial para demonstrar que o provedor está agindo em conformidade com a legislação de direitos autorais, evitando o uso indevido de materiais protegidos.



- **Controle Interno e Procedimentos para Resolução de Disputas:**

Revisões Regulares de Conformidade: O Código recomenda que os provedores estabeleçam procedimentos internos para revisar regularmente a conformidade com direitos autorais, assegurando que todos os dados e conteúdos utilizados nos modelos de IA estejam devidamente licenciados ou que se enquadrem nas exceções legais aplicáveis.

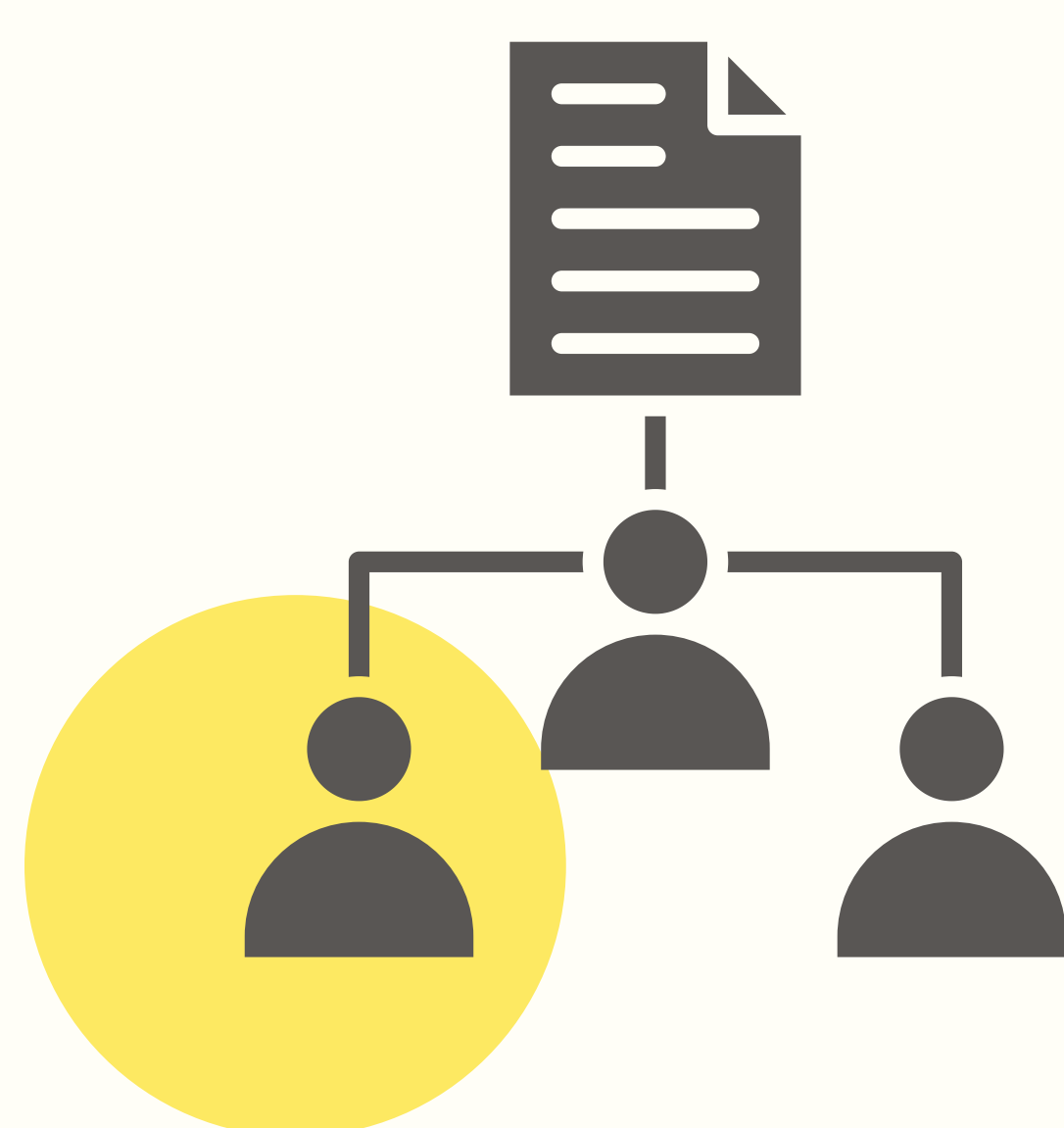
Procedimentos de Resolução de Disputas: Caso ocorra uma reivindicação de direitos autorais ou disputa envolvendo conteúdo utilizado, os provedores devem ter processos definidos para responder a essas demandas. Isso pode incluir a remoção do conteúdo ou a revisão do modelo para eliminar o uso indevido.

- **Monitoramento e Controle de Conteúdos Gerados:**

Rastreamento de Uso de Dados Protegidos: Provedores devem implementar mecanismos que permitam rastrear a origem dos dados utilizados e controlar se os dados protegidos estão sendo usados conforme o permitido. Isso inclui a aplicação de ferramentas para monitorar o uso do conteúdo e alertar caso o modelo esteja reproduzindo material de maneira que infrinja os direitos autorais.

Supervisão de Conteúdos Derivados: O Código sugere que os provedores monitoram se o conteúdo gerado pelo modelo pode implicar uma derivação de obras protegidas. O objetivo é garantir que a IA não produza conteúdos que copiem ou reproduzam diretamente obras protegidas por direitos autorais, o que poderia resultar em violações involuntárias.

2. Taxonomia de Riscos Sistêmicos



Esse item ajuda a identificar e categorizar os riscos que um modelo de IA pode apresentar em uma escala maior, especialmente se tiver impacto em sistemas críticos ou na sociedade.

• Tipos de Riscos Sistêmicos

Riscos associados a sistemas de IA que poderiam afetar grandes grupos ou setores. Inclui:

Segurança Cibernética (Artigo 3(65)): IA que possa descobrir ou explorar falhas de segurança, facilitando ataques cibernéticos.

CBRN (Químico, Biológico, Radiológico e Nuclear): Capacidades que poderiam ser usadas para desenvolver armas ou tecnologias perigosas.

Manipulação em Massa: A IA pode facilitar campanhas de desinformação ou manipular a opinião pública, o que afeta a democracia.

Discriminação em Grande Escala: Riscos de a IA discriminar pessoas ou grupos com base em dados enviesados.

• Características dos Riscos

São características que ajudam a avaliar a gravidade dos riscos, incluindo:

Origem: Se o risco vem das capacidades do modelo ou de como ele é distribuído.

Intenção: Se o risco é intencional ou acidental.

Probabilidade e Severidade: A chance de o risco se concretizar e o quão grave seria seu impacto.

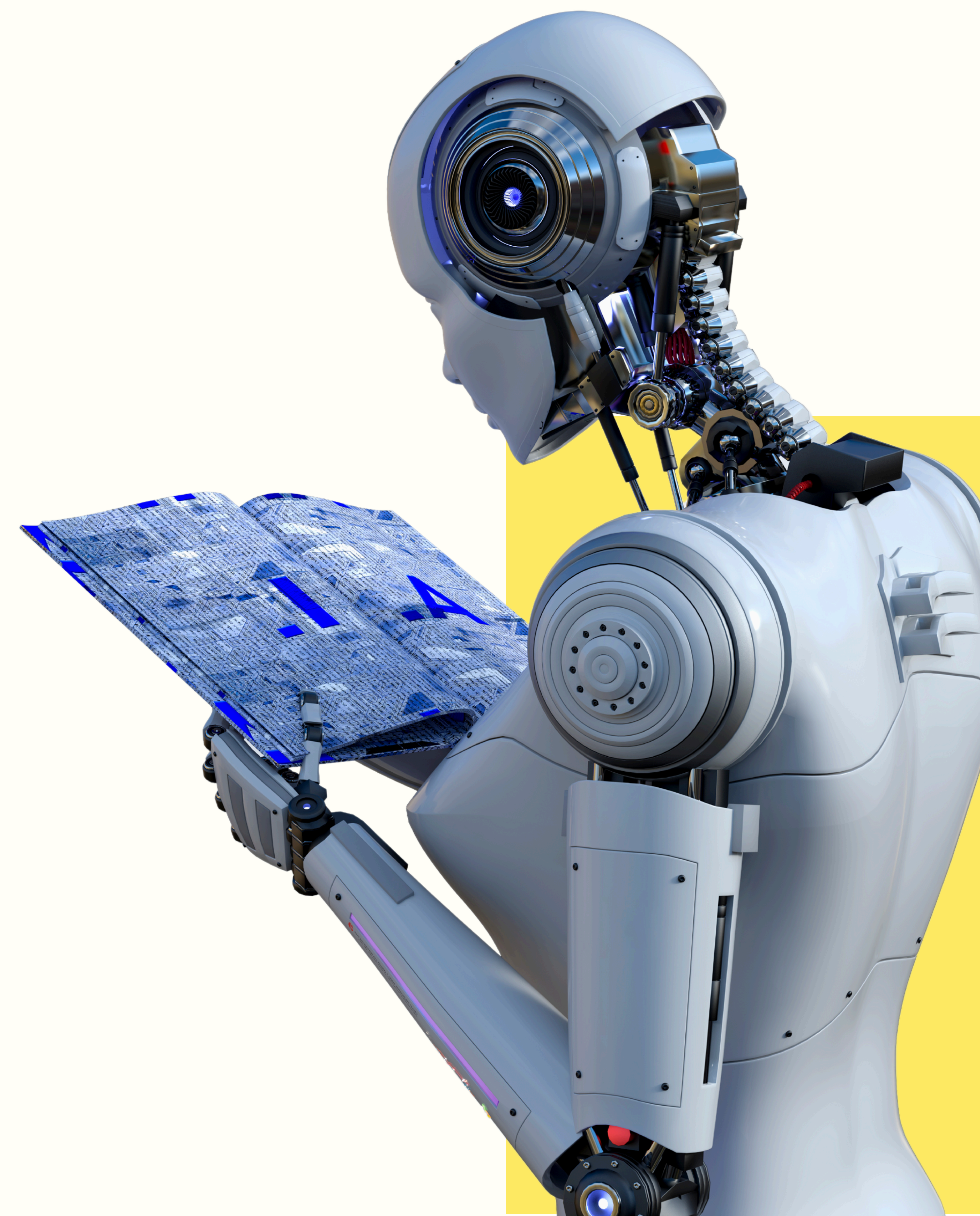
- **Fontes de Riscos**

Identifica o que pode causar riscos, como:

Capacidades Perigosas: Se o modelo tem capacidades de ofensiva cibernética, autonomia ou manipulação.

Propensões do Modelo: Se o modelo tende a agir fora do controle humano ou apresenta viés.

Contexto e Configurações: Fatores externos, como acesso a dados ou ferramentas, que podem aumentar os riscos.



3) Regras para IA com Risco Sistêmico

Essas são regras especiais para IA que apresenta riscos significativos, como aquelas que poderiam afetar setores inteiros ou causar danos significativos.



- **Estrutura de Segurança e Mitigação de Riscos (Artigo 55(1))**

Identificação de Riscos: O provedor precisa identificar e especificar os principais riscos associados ao modelo e como eles podem se manifestar.

Análise e Coleta de Evidências: O provedor deve realizar análises contínuas para entender como o modelo pode causar esses riscos e deve coletar evidências durante toda a vida do modelo (antes, durante e depois do treinamento).

- **Ciclo de Vida da Avaliação de Risco**

Pré-Treinamento: Antes de começar o treinamento, o provedor deve garantir que todas as práticas de segurança estão em vigor.

Durante o Treinamento: Reavaliar os riscos periodicamente durante o treinamento.

Pós-Implantação: Após a implantação, o provedor deve monitorar os riscos e verificar continuamente se o modelo ainda está seguro para uso.

4) Mitigação Técnica de Risco

Medidas para reduzir riscos associados ao uso da IA, especialmente em áreas de segurança e proteção de dados.

- **Medidas de Proteção e Segurança Cibernética:**

Mitigação de Segurança: Definir medidas para proteger o modelo, como modificar seu comportamento ou implementar salvaguardas para impedir que ele seja usado de forma perigosa.

Proteção Cibernética (Artigo 55(1)(d)): Medidas para proteger o modelo contra vazamento de dados ou acesso não autorizado. Pode incluir criptografia, controles de acesso, e monitoramento.



- **Relatório de Segurança e Segurança (SSR):**

Relatório detalhado com todas as avaliações de risco e mitigação feitas para o modelo. Esse relatório ajuda a orientar decisões sobre continuar ou parar o desenvolvimento e implantação do modelo.

5) Governança de Riscos Sistêmicos

Regras para assegurar que todos os níveis da organização (incluindo diretoria e gerência) estejam envolvidos na segurança e mitigação de riscos, especialmente para provedores de modelos de IA com risco sistêmico.



- **Responsabilidade Organizacional (Artigo 55(1)):**

Nível Executivo e Diretoria: Exige-se que as organizações aloquem responsabilidades para a alta gestão e diretoria, com o objetivo de acompanhar e mitigar os riscos sistêmicos que os modelos possam representar. Isso pode incluir a criação de um comitê de risco para monitorar e avaliar continuamente a adequação das práticas de segurança do modelo.

- **Avaliação Periódica de Conformidade e Adequação (Artigo 55(1) e Recital 114):**

Revisão Anual: as organizações façam uma revisão anual para verificar a conformidade e adequação das medidas de segurança e mitigação de riscos. Essa avaliação deve considerar atividades planejadas e mudanças no contexto tecnológico e de ameaças.

- **Avaliação Independente de Riscos e Mitigações (Artigo 55(1)):**

Antes e Após a Implantação: Requer que avaliações de risco independentes sejam realizadas antes e após a implantação do modelo, proporcionando uma validação externa das práticas de mitigação e segurança adotadas. Isso pode envolver especialistas independentes ou reguladores que verifiquem a eficácia das medidas de mitigação aplicadas ao longo da vida útil do modelo.

Considerações Finais e Adaptações para o Futuro

O documento final, previsto para 02/05/2025, consolidará essas diretrizes e representará um compromisso robusto para o uso responsável da IA de propósito geral na União Europeia, em conformidade com o EU AI Act, e pode servir como modelo global de boas práticas, inclusive para o Brasil.

Autores



Rony Vainzof
Sócio
rony@vlklaw.com.br



Mateus Lamonica
Estagiário
mateus.lamonica@vlklaw.com.br