IA Generativa

No Serviço Público

Definições, usos e boas práticas

Serpro

MINISTÉRIO DA FAZENDA

MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS

GOVERNO FEDERAL

BRASIL

UNIÃO E RECONSTRUÇÃO

MINISTÉRIO DA GESTÃO E DA INOVAÇÃO EM SERVIÇOS PÚBLICOS

Esther Dweck
Ministra da Gestão e da Inovação em Serviços Públicos

SECRETARIA DE GOVERNO DIGITAL

Rogério Souza Mascarenhas
Secretário de Governo Digital

DIRETORIA DE INFRAESTRUTURA DE DADOS

Renan Mendes Gaya Lopes Dos Santos
Diretor de Infraestrutura de Dados

COORDENACAO-GERAL DE FOMENTO DA INTELIGÊNCIA ARTIFICIAL

Thaciana Guimarães De Oliveira Cerqueira
Coordenadora-Geral de Fomento da Inteligência Artificial Responsável

EQUIPE RESPONSÁVEL PELA ELABORAÇÃO

Camila Chagas Patriota
Carlos Ely Pimenta
Carlos Rodrigo Fonseca Lima
Denise Rodrigues de Araújo Merschmann
Flábia Alves Lopes
Italo Alberto do Nascimento Sousa
Juliana De Freitas Ulisses Machado
Marina Geralda de Aguiar
Rohgi Toshio Meneses Chikushi
Welsinner Gomes de Brito

Introdução ao Guia

Este guia foi elaborado pela Secretaria de Governo Digital (SGD) e pelo Serviço Federal de Processamento de Dados (SERPRO), com a participação dos demais membros do Núcleo de Inteligência Artificial (IA) do Governo, como a Casa Civil, MCTI, ENAP e Dataprev, com o propósito de apoiar servidores públicos no entendimento e uso responsável de ferramentas de Inteligência Artificial Generativa (IAG).

A ação faz parte do Plano Brasileiro de Inteligência Artificial e visa promover o uso consciente e ético da IA no setor público. Neste guia você encontrará uma visão geral da IAG, explicações claras sobre seus desafios e limitações, além de orientações práticas para o uso ético e eficaz dessas ferramentas no setor público brasileiro.

Nosso objetivo é orientar para o uso responsável de IAG, a fim de otimizar processos, tomar decisões mais eficazes e aprimorar o atendimento ao cidadão em conformidade com os princípios éticos da IA e rigor técnico em sua aplicação.

Esta cartilha ajudará a navegar pelas ferramentas de IAG, maximizando seus benefícios e mitigando os riscos, e será atualizada continuamente para refletir as mudanças nas regulamentações e avanços tecnológicos, garantindo que você esteja sempre bem informado.

Inteligência Artificial Generativa

A Inteligência Artificial Generativa é um ramo da inteligência artificial projetado para criar conteúdos, como textos, áudios, imagens, vídeos e até códigos de software. Essa tecnologia se baseia em modelos computacionais avançados, que aprendem padrões a partir de grandes volumes de dados.

No **setor público**, as ferramentas de IAG oferecem oportunidades para modernizar processos, como automação de tarefas repetitivas, redação de documentos e criação de materiais visuais. Contudo, sua adoção deve ser feita de forma criteriosa, garantindo que os benefícios superem os riscos e que o uso esteja alinhado com valores éticos e regulatórios.

Fases de Desenvolvimento

Modelos de inteligência artificial normalmente passam por ciclos compostos por duas grandes fases: **treinamento** – onde ocorre o ajuste e captura de padrões nos dados, e **produção** - fase da aplicação.

Durante a fase de treinamento, esses modelos são expostos a dados diversos, como documentos, imagens e outras fontes públicas ou privadas, para identificar relações e estruturas. Assim, conseguem gerar conteúdos que parecem ter sido criados por humanos. Por exemplo, podem criar um relatório, compor um texto ou gerar um design com base em comandos específicos fornecidos por usuários.

Outra característica importante é que a IAG funciona com base em probabilidade e padrões, e não possui "consciência" ou "intenção". Isso significa que, embora seja uma ferramenta poderosa, ela pode gerar respostas que, em alguns casos, não são adequadas ou estão fora de contexto, o que exige atenção e validação por parte dos usuários.

É crucial entender que, como esses modelos dependem dos dados com os quais foram treinados, eles podem reproduzir vieses ou informações imprecisas presentes nesses dados. Isso reforça a necessidade de supervisão humana e de uma avaliação cuidadosa de suas aplicações, especialmente em instituições públicas que lidam com informações sensíveis e de impacto social.

Modelos de IA Generativa

Os modelos de IAG podem ser classificados por licenças **abertas** ou **fechadas**. A diferença entre eles está diretamente relacionada à forma como podem ser utilizados e adaptados às necessidades específicas de cada contexto.

Enquanto os modelos fechados são projetados para oferecer soluções padronizadas e práticas, os modelos abertos se destacam pela flexibilidade, permitindo personalizações e maior controle por parte do usuário. A escolha entre eles depende das necessidades do projeto, dos recursos disponíveis e das preocupações éticas envolvidas.

Em termos de desempenho, os modelos fechados frequentemente oferecem resultados otimizados em tarefas gerais devido ao alto investimento das empresas de tecnologia em infraestrutura e treinamento contínuo. Já os modelos abertos, embora demandem mais esforço para ajustes, podem superar os fechados em casos de uso específico, especialmente quando adaptados com dados internos.

A escolha entre os dois tipos de modelos deve equilibrar a disponibilidade de recursos, eficiência, privacidade, custo e a necessidade de personalização, garantindo que as ferramentas de IAG sejam usadas de forma responsável e alinhada aos objetivos institucionais.

Modelos Fechados

Desenvolvidos principalmente por grandes empresas como OpenAI, Google e Anthropic, são amplamente utilizados para aplicações como assistentes em bases de conhecimento e bases de dados, auxílio na geração de relatórios padronizados, e automação de atendimento.

Esses modelos geralmente são acessados por meio de APIs, o que facilita sua integração em sistemas existentes e permite que as instituições utilizem suas funcionalidades sem a necessidade de infraestrutura complexa.

Você Sabia?

No mercado brasileiros, para modelos fechados destaca-se a Maritaca IA, que fornece um modelo treinado, utilizando bases de língua portuguesa, e que pode atender às demandas locais com um custo reduzido e performance equivalente a essas outras soluções.

Atualmente, os modelos dessas empresas que oferecem os melhores benchmarks e desempenhos são: o GPT-4 e o 3-mini da OpenAI, o Gemini do Google, e o Claude da Anthropic. Apesar das vantagens em desempenho e da facilidade no uso, os modelos fechados apresentam desafios importantes. Muitas vezes, seu uso implica o envio de dados para servidores localizados no exterior, o que pode representar riscos à privacidade e à conformidade com regulamentos locais.

Além disso, esses modelos funcionam como “caixas pretas”, dificultando auditorias e adaptações, o que pode ser uma limitação para instituições públicas que lidam com informações sensíveis ou que precisam de maior transparência.

Modelos Abertos

Desenvolvidos principalmente por empresas como a Meta, com o LLaMA 3, pela europeia Mistral AI, com o Mistral 7B, e recentemente pelo DeepSeek, com o DeepSeek-R1, esses modelos oferecem vantagens significativas para quem busca maior controle e flexibilidade.

Esses modelos são disponibilizados sob licenças abertas, permitindo que sejam utilizados e ajustados para atender a contextos específicos e podem ser instalados e executados em servidores locais, eliminando a necessidade de enviar dados para terceiros e garantindo a segurança de informações sensíveis.

A Hugging Face se tornou a principal plataforma para acesso a esses modelos abertos. Com suporte dessa plataforma, os usuários podem acessar esses modelos, fazer instalações locais, e realizar customizações tanto locais como diretamente na plata-

No Brasil, a Maritaca AI destaca-se com o Sabia-7B, voltado para o idioma português, demonstrando performance equivalente a um custo reduzido.

forma, permitindo que os usuários possam utilizar esses modelos de modo simplificado sem necessidade de treinamento do zero. Esses modelos são disponibilizados sob licenças abertas e podem ser ajustados para atender a contextos específicos, sendo executados em servidores locais para garantir maior segurança e controle sobre os dados.

A plataforma Hugging Face oferece também mecanismos para download, fine-tuning e inferência na nuvem de forma simplificada, permitindo que os órgãos possam utilizar esses modelos pré-treinados sem precisar de infraestrutura própria.

Quais as tendências?

Como tendência na área da IAG podemos destacar a **Geração Aumentada via Recuperação** - RAG (Retrieval-Augmented Generation). Trata-se de uma das inovações mais promissoras para lidar com gestão de conhecimento nesse contexto. RAG é uma técnica que combina síntese de texto com recuperação de informação, integrando uma base de conhecimento ao modelo de IAG. Isso permite respostas mais precisas e contextualmente aderentes, ao complementar a capacidade do LLM com dados mais relevantes.

Com o RAG, o modelo não é ajustado diretamente, mas combinado a sistemas de informações que fornecem bases de dados que podem ser consultadas para a geração de respostas. Embora os modelos fechados sejam projetados para tarefas generalistas, eles também podem ser integrados a dados próprios de uma instituição por meio de técnicas como RAG.

Princípios Fundamentais para o Uso da IAG

Legalidade

O uso da IAG deve estar em conformidade com as leis vigentes, incluindo a Lei Geral de Proteção de Dados (LGPD) e demais normativos, garantindo que os direitos dos cidadãos sejam respeitados.

Impessoalidade

Decisões assistidas por IAG devem ser imparciais, baseadas em critérios objetivos e livres de viés ou preconceito.

Moralidade

Essencial no uso de IAG, assegurando transparência, justiça e respeito aos direitos humanos e às garantias democráticas.

Publicidade

Os processos envolvendo IAG devem ser transparentes, permitindo auditoria e contestação quando necessário.

Eficiência

Ferramentas de IAG devem ser usadas para automatizar tarefas repetitivas, liberando recursos para atividades mais estratégicas, desde que sempre com revisão humana.

Riscos

Antes de utilizar ferramentas de IAG em processos de trabalho, é necessário analisar os benefícios e riscos associados a cada caso de uso. A identificação das vulnerabilidades específicas e a implementação de medidas de mitigação são passos indispensáveis para garantir que as ferramentas sejam usadas de forma ética e segura. O uso de IAG está sujeito a riscos inerentes, como:

1. Alucinações na IA Generativa

Modelos de IAG são treinados em grandes volumes de dados e podem produzir respostas "criativas" baseadas em padrões e informações aprendidas durante um longo processo de treinamento. Entretanto, essa criatividade sintética pode desencadear **alucinações** – situações em que um modelo de inteligência artificial gera informações inexatas, irrelevantes ou totalmente fabricadas, mesmo que pareçam confiáveis e coerentes. Esse comportamento surge devido a limitações nos dados de treinamento, dados de entrada e na arquitetura do modelo, tais como:

Perguntas mal formuladas

Questões ambíguas ou abertas podem levar a respostas imprecisas.

Dados incompletos ou enviesados

O modelo pode basear sua resposta em informações limitadas ou distorcidas presentes no conjunto de dados.

Tendência a completar informações

Quando não encontra uma resposta clara, o modelo pode "criar" com base em dados relacionados.

Falta de atualização

Os modelos são treinados com dados disponíveis até uma determinada data e, após isso, não conseguem incluir automaticamente informações mais recentes. Isso significa que, para assuntos contemporâneos ou em constante evolução, as respostas podem estar desatualizadas.

2. Direito Autoral e Propriedade Intelectual

É essencial que o uso de ferramentas de IAG seja acompanhado de uma avaliação rigorosa sobre a origem e a natureza dos dados gerados, garantindo que estejam em conformidade com a legislação e as boas práticas institucionais. Este item aborda os cuidados necessários em relação ao uso de ferramentas de IAG, considerando aspectos legais e éticos ligados a direitos autorais e propriedade intelectual.

A. Dados Protegidos

Muitas ferramentas de IAG são treinadas com dados que podem estar protegidos por direitos autorais e propriedade intelectual. Embora a coleta e o tratamento desses dados possam ser legais, nem sempre isso é garantido.

B. Obras Protegidas

Trechos de textos, imagens ou outros conteúdos gerados pela IAG podem conter elementos derivados de obras protegidas, expondo o órgão a riscos jurídicos e reputacionais.

3. Vazamento de Dados, Ataques e Acessos Não Autorizados

O uso de modelos de inteligência artificial generativa pode expor organizações e usuários a uma série de riscos relacionados à segurança da informação, caso não sejam adotadas práticas robustas de mitigação.

A. Vazamento de Dados

Modelos generativos treinados em grandes volumes de dados podem, inadvertidamente, restituir informações sensíveis presentes no conjunto de dados original, especialmente se os dados não foram devidamente anonimizados. Além disso, práticas inadequadas de armazenamento e compartilhamento podem expor informações pessoais ou confidenciais.

4. Importância da Anonimização de Dados

A anonimização de dados é uma prática essencial para garantir que as informações pessoais e sensíveis não sejam vinculadas a indivíduos ou entidades específicas, reduzindo o risco de exposição acidental. Técnicas como mascaramento, tokenização, generalização e substituição de dados reais por sintéticos são recomendadas para proteger a privacidade dos indivíduos.

5. Viés na IAG

Apresentado como uma questão central devido à dependência dos modelos de IAG de grandes volumes de dados de treinamento. Esses dados podem conter preconceitos implícitos, desigualdades ou representações inadequadas, que acabam sendo refletidos nos resultados gerados.

A. Viés de Contexto

Ocorre quando a IA falha em interpretar ou aplicar os dados corretamente em contextos diferentes daquele em que foi treinada.

B. Viés de Automação

É a tendência de confiar nas decisões automatizadas da IA como sendo sempre precisas ou imparciais, mesmo quando estão incorretas.

C. Viés de Representatividade

Dados de treinamento que não representam adequadamente a diversidade podem fazer com que o modelo funcione bem para alguns grupos, mas mal para outros.

D. Viés de Seleção

Ocorre quando os dados usados no treinamento não são suficientemente aleatórios ou representativos, resultando em uma amostra enviesada.

E. Viés de Exclusão

Grupos sub-representados ou excluídos do conjunto de dados acabam sendo ignorados ou mal interpretados pela IA.

Classificação de Riscos

O Projeto de Lei nº 2338/2023 em discussão no Congresso Nacional busca estabelecer um marco regulatório para o desenvolvimento e uso de sistemas de inteligência artificial (IA) no Brasil. A proposta estabelece uma classificação dos riscos da IA, adequando as exigências conforme o potencial impacto dos sistemas.

Com base nesse projeto, esta cartilha adota uma categorização que define riscos excessivos, altos riscos e demais riscos, determinando o regime jurídico aplicável e as obrigações correspondentes para garantir a segurança, transparência e proteção dos direitos fundamentais.

1. Risco Excessivo

O PL 2338/2023 propõe a proibição do desenvolvimento e uso de sistemas de inteligência artificial que apresentem riscos excessivos à sociedade e aos direitos fundamentais. São considerados de risco excessivo e vedados os sistemas de IA que:

A. Induzam ou manipulem comportamentos de forma prejudicial à saúde, segurança ou direitos fundamentais.

B. Explorem vulnerabilidades de grupos vulneráveis (ex.: crianças, idosos, pessoas com deficiência) para influenciar ou causar danos.

C. Realizem perfis preditivos criminais, avaliando características pessoais para prever crimes ou reincidência.

D. Sejam utilizados pelo poder público para pontuação social, classificando cidadãos com base em seu comportamento ou personalidade de maneira discriminatória.

E. Façam identificação biométrica em tempo real em espaços públicos, salvo exceções (ex.: busca de pessoas desaparecidas).

2. Alto Risco

Apesar desses sistemas não serem proibidos, o seu uso deve ser cuidadosamente monitorado e controlado para evitar abusos, discriminações ou danos irreversíveis, podendo gerar impactos significativos sobre direitos fundamentais, segurança, privacidade e bem-estar da sociedade.

Dessa forma, a categorização de alto risco impõe obrigações regulatórias mais rígidas e medidas específicas de governança, exigindo avaliação de impacto algorítmico, transparência, supervisão humana e mitigação de riscos, garantindo que a IA seja usada de forma ética e responsável.

São considerados de alto risco os sistemas de IA utilizados em:

1. Em segurança para gestão de infraestruturas críticas, como trânsito e redes de abastecimento.

2. Aplicações na educação e formação profissional, incluindo acesso e avaliação de estudantes.

3. Uso em recrutamento, triagem, avaliação de candidatos, gestão e monitoramento no trabalho.

4. Avaliação de elegibilidade e concessão de serviços públicos e privados essenciais.

5. Análise da capacidade de endividamento e classificação de crédito.

6. Definição de prioridades para serviços de emergência, como bombeiros e assistência médica.

7. Apoio à administração da justiça, investigação e aplicação da lei.

8. Uso de veículos autônomos com riscos à integridade física.

9. Aplicações na saúde, incluindo diagnósticos e procedimentos médicos.

10. Uso de sistemas biométricos de identificação.

11. Investigação criminal e segurança pública, incluindo avaliação de riscos e perfis criminais.

12. Análise de crimes por meio de grandes conjuntos de dados para identificar padrões.

13. Investigação administrativa para avaliar provas e prever infrações com base em perfis.

14. Gestão da migração e controle de fronteiras.

3. Risco Moderado

Sistemas que apresentam um impacto significativo, mas não tão elevado quanto os sistemas classificados como de alto risco. Esses sistemas podem afetar a vida das pessoas de maneira relevante, porém o potencial de dano é limitado ou controlável, não sendo considerado grave o suficiente para justificar um nível de regulamentação tão rigoroso quanto os de risco alto ou excessivo.

Entretanto, exige-se a implementação de medidas de transparência, explicabilidade e supervisão, mas com menos rigidez do que os sistemas de alto risco.

Exemplos de sistemas de risco moderado incluem:

1. IA em áreas como recursos humanos, por exemplo, para triagem inicial de currículos ou entrevistas automatizadas, onde as decisões são importantes, mas não definitivas.

2. Sistemas de recomendação em comércio eletrônico, que influenciam o consumo, mas sem implicações diretas na saúde, liberdade ou outros direitos fundamentais dos indivíduos.

3. Análise de dados em setores como marketing ou publicidade, onde a IA é usada para personalizar ofertas, mas sem invadir a privacidade de forma invasiva.

4. Baixo Risco

Esses sistemas não exigem medidas rigorosas de regulação, mas ainda devem cumprir requisitos básicos de transparência, responsabilidade e boas práticas de governança. Embora esses sistemas tenham baixo potencial de dano, a lei ainda exige que sejam transparentes quanto ao seu funcionamento, especialmente quando interagem diretamente com pessoas, para garantir que os usuários saibam quando estão lidando com IA e possam tomar decisões informadas.

Exemplos de sistemas de baixo risco incluem:

1. Chatbots e assistentes virtuais, que interagem com usuários sem tomar decisões com consequências jurídicas ou econômicas relevantes.

2. Filtros de recomendação, como os usados em plataformas de streaming e comércio eletrônico, para sugerir conteúdos ou produtos com base em preferências do usuário.

3. Ferramentas de edição de imagem e vídeo, que modificam conteúdos visuais sem comprometer direitos fundamentais.

4. Jogos e aplicações recreativas, que utilizam IA para melhorar a experiência do usuário sem riscos significativos.

Recomendações e Boas Práticas

Ao lidar com IAG, é necessário adotar uma abordagem consciente para evitar riscos como o vazamento de dados sensíveis, disseminação de informações enganosas e interpretações equivocadas de resultados. Isso reflete um compromisso com a ética, a privacidade e a segurança, especialmente no contexto da administração pública.

Para promover uma cultura de responsabilidade no uso dessas ferramentas, algumas regras básicas devem ser seguidas.

1. Evitar informações sensíveis

Ao compartilhar informações sensíveis ou dados pessoais, lembre-se que a segurança desses dados depende de vários fatores. Consulte as políticas de privacidade de dados sensíveis e pessoais e cuidado ao compartilhar tais informações.

2. Os três "Como" a considerar

Como as perguntas serão usadas?

Evite inserir informações que você não compartilharia publicamente, considerando que a ferramenta aprende com os dados fornecidos.

Como as respostas podem ser enganosas?

As respostas da IA, embora pareçam confiáveis, podem ser imprecisas ou derivadas de fontes não verificadas, exigindo validação rigorosa.

Como a IA funciona?

As respostas são geradas probabilisticamente, sem compreensão de contexto ou viés, necessitando interpretação criteriosa.

3. Consultas especializadas

Servidores públicos devem buscar orientação de serviços jurídicos e especialistas em privacidade e segurança de dados e informações para decidir sobre o uso adequado dessas ferramentas.

A partir dessas regras iniciais, as seguintes orientações devem ser consideradas em relação a boas práticas

1 O uso de IAG deve passar por revisão institucional, considerando os princípios éticos, os riscos e as melhores práticas definidos internamente pelo órgão.

A Revisar o conteúdo gerado para garantir alinhamento com princípios institucionais. Analise a resposta dos prompts de IAG para garantir que eles atendam os padrões da Instituição quanto aos princípios éticos, de legalidade, e adequação.

BOM USO

Rever a resposta de IAG e ajustar para alinhamento com os valores institucionais antes da publicação.

NÃO RECOMENDÁVEL

Publicar diretamente conteúdos gerados, sem revisão, especialmente em temas sensíveis.

B Sempre avalie criteriosamente e revise o conteúdo gerado por IAG, ainda que a resposta pareça confiável.

BOM USO

Testar a IAG para gerar rascunhos e verificar cada ponto factual antes de aceitar o conteúdo respondido.

NÃO RECOMENDÁVEL

Aceitar respostas de IAG como verdadeiras sem verificá-las, principalmente em relação a tópicos técnicos ou científicos envolvendo operações matemáticas.

2 Decisões automatizadas pela IAG não devem ser adotadas sem revisão humana, principalmente quando usada para informar o público externo ou quando envolver decisões estratégicas.

A Evite o uso sem validação e considere a capacidade de identificar imprecisões antes de usar a IAG.

BOM USO

Solicitar apoio de especialistas na validação de conteúdos gerados por IA antes de sua publicação.

NÃO RECOMENDÁVEL

Utilizar conteúdos técnicos complexos gerados por IA sem consultar especialistas na área.

3 O uso de IAG deve estar alinhado com o código de conduta institucional e as políticas de não discriminação, garantindo que o conteúdo seja apropriado e não discriminatório.

A Avaliar o conteúdo da resposta da IA para garantir que não discrimine indivíduos com base em raça, cor, religião, sexo, nacionalidade, idade, deficiência, estado civil, afiliação política ou orientação sexual.

BOM USO

Usar IA para criar descrições inclusivas, validando que a linguagem não perpetue estereótipos.

NÃO RECOMENDÁVEL

Utilizar conteúdo sem avaliação, resultando em mensagens discriminatórias ou tendenciosas.

4

Para proteger servidores, cidadãos e a instituição, recomenda-se evitar o uso de conteúdo criado pela IA que seja inapropriado, discriminatório, incorreto ou prejudicial.

A

Revisar o conteúdo gerado para garantir alinhamento com princípios institucionais. Analise a resposta dos prompts de IAG para garantir que eles atendam os padrões da Instituição quanto aos princípios éticos, de legalidade, e adequação.

BOM USO

Rever a resposta de IA e ajustar para alinhamento com os valores institucionais antes da publicação.

NÃO RECOMENDÁVEL

Publicar diretamente conteúdos gerados, sem revisão, especialmente em temas sensíveis.

B Avaliar o conteúdo da resposta da IA para garantir que não discrimine indivíduos com base em raça, cor, religião, sexo, nacionalidade, idade, deficiência, estado civil, afiliação política ou orientação sexual.

5 A responsabilidade do servidor sobre qualquer documento produzido, com ou sem IAG, permanece inalterada. O uso inadequado de IAG não exime o servidor de revisar e assumir a autoria plena do resultado.

A Identificar a resposta gerada pela IA, sem prejuízo da responsabilização do servidor.

BOM USO

Incluir no rodapé de um documento institucional a observação: *"Parte do conteúdo foi gerado com o auxílio de IA."*

NÃO RECOMENDÁVEL

Apresentar conteúdo gerado por IA como criação exclusiva de autores humanos, sem a devida transparência.

6 Recomenda-se evitar o uso de endereços de e-mail, credenciais e números de telefone do órgão ou entidade para criar contas em plataformas externas de IAG, prevenindo a ligação entre uso pessoal e trabalho institucional.

A Não use credenciais da instituição, endereços de e-mail ou números de telefone como login para aplicativos de IAG disponíveis publicamente.

BOM USO

Criar contas usando credenciais genéricas, desvinculadas da instituição, para experimentação em IAG.

NÃO RECOMENDÁVEL

Usar o e-mail institucional para acessar um serviço de IA, como o ChatGPT ou qualquer prompt semelhante, sem a devida aprovação.

7 Para proteger informações sensíveis e confidenciais, incluindo dados protegidos por lei e propriedade intelectual, é indicado que servidores e prestadores de serviços utilizem apenas soluções de IAG aprovadas pelo órgão.

A Não insira informações internas da instituição em aplicativos de IAG que não seja uma solução aprovada.

BOM USO

Solicitar sugestões de estratégias de gerenciamento genéricas, sem mencionar nomes de projetos ou informações específicas da instituição.

NÃO RECOMENDÁVEL

Compartilhar detalhes de contratos, projetos em andamento ou políticas internas ao solicitar sugestões.

B

Não insira informações pessoais de servidores, cidadãos ou terceiros em aplicativos de IAG não aprovados.

BOM USO

Usar informações fictícias ou simuladas para criar exemplos ou cenários durante testes com IA.

NÃO RECOMENDÁVEL

Fornecer dados reais de servidores, como CPF ou informações de saúde, para obtenção de insights.

C

Garanta que os conteúdos gerados respeitem direitos autorais e de propriedade intelectual.

BOM USO

Solicitar à IA conteúdo genérico e verificar sua originalidade antes de incorporá-lo a materiais oficiais.

NÃO RECOMENDÁVEL

Usar trechos gerados por IA sem verificar se são plágios ou violações de direitos autorais.

Considerações Finais

Este texto preliminar estabelece os fundamentos e princípios gerais para as boas práticas do uso da Inteligência Artificial generativa no âmbito da administração pública federal. Os pontos visam proteger contra vazamentos de dados, garantir a confidencialidade das informações sigilosas, evitar violações de propriedade intelectual, assegurar a segurança de servidores e cidadãos, prevenir danos à reputação da instituição, mitigar viés de modelos automatizados e evitar infrações de direitos autorais.

Referências

1. AUSTRALIAN GOVERNMENT ARCHITECTURE. Interim guidance on government use of public generative AI tools. November 2023. Disponível em: https://architecture.digital.gov.au/guidance-generative-ai

Acesso em: 31 jan. 2025.

2. BRASIL. Receita Federal. Instrução Normativa RFB nº 2.134, de 10 de março de 2023. Disponível em: http://normas.receita.fazenda.gov.br/sijut2consulta/link.action?idAto=138693.

Acesso em: 31 jan. 2025.

3. BRASIL. Projeto de Lei n° 2338, de 2023. Autoria: Senador Rodrigo Pacheco (PSD/MG). Disponível em: https://www25.senado.leg.br/web/atividade/materias/-/materia/157233.

Acesso em: 31 jan. 2025.

4. BRASIL. Autoridade Nacional de Proteção de Dados. Radar Tecnológico: Inteligência Artificial Generativa. Disponível em: https://www.gov.br/anpd/pt-br/centrais-de-conteudo/documentos-tecnicos-orientativos/radar_tecnologico_ia_generativa_anpd.pdf.

Acesso em: 31 jan. 2025.

5. BRASIL. Tribunal de Contas da União. Guia de uso de inteligência artificial generativa no Tribunal de Contas da União (TCU). Disponível em: https://portal.tcu.gov.br/data/files/42/F7/91/4B/B59019105E366F09E18818A8/Guia%20de%20uso%20de%20IA%20generativa%20no%20TCU.pdf.

Acesso em: 31 jan. 2025.

6. BRASIL. Controladoria-Geral da União (CGU). Guia de Uso Responsável de Inteligência Artificial. Brasília: CGU, 2024.
Disponível em: https://repositorio.cgu.gov.br/bitstream/1/94244/2/Guia_de_uso_responsavel_de_IA_v5_publicacao.pdf.

Acesso em: 31 jan. 2025.

7. CANADA. Guide on the Use of Generative AI.
Disponível em: https://www.canada.ca/en/government/system/digital-government/digitalgovernment-innovations/responsible-use-ai/guide-use-generative-ai.html.

Acesso em: 31 jan. 2025.

8. UNITED KINGDOM. Guidance to civil servants on use of generative AI.
Disponível em: https://www.gov.uk/government/publications/guidance-to-civil-servants-on-use-of-generative-ai/guidance-to-civil-servants-on-use-of-generative-ai.

Acesso em: 31 jan. 2025.

Imagens obtidas a partir de banco de imagens gratuito : Freepik

MINISTÉRIO DA
FAZENDA

MINISTÉRIO DA
GESTÃO E DA INOVAÇÃO
EM SERVIÇOS PÚBLICOS

GOVERNO FEDERAL

BRASIL

UNIÃO E RECONSTRUÇÃO

gov.br

Serpro