



**PREFEITURA DE
SÃO PAULO**

MANUAL DE PRÁTICAS E AÇÕES PARA IMPLEMENTAÇÃO DA POLÍTICA DE INTELIGÊNCIA ARTIFICIAL (IA)

1ª EDIÇÃO - 2026



**PREFEITURA DE
SÃO PAULO**

PREFEITURA DE SÃO PAULO

Prefeito
Ricardo Nunes

**SECRETARIA MUNICIPAL
DE GESTÃO**

Secretária
Marcela Arruda

Secretária Adjunta
Regina Silverio

Chefe de Gabinete
Thaís Rodrigues

**COORDENADOR DE
TECNOLOGIA DA INFORMAÇÃO
E COMUNICAÇÃO**
Wagner Santana Silveira

REDAÇÃO
Davi Giugno
João Victor Silva
Rafael da Matta

DIAGRAMAÇÃO
Bianca Abreu

REVISÃO
Roseli Andrade Pires

Este manual tem como objetivo orientar a implementação prática da Política de Inteligência Artificial (IA) no âmbito da Secretaria Municipal de Gestão (SEGES), detalhando ações, fluxos e responsabilidades para garantir a aplicação efetiva dos princípios e diretrizes estabelecidos na Portaria correspondente.

Iniciaremos com um glossário dos termos comumente utilizados no contexto da IA, e a seguir estabeleceremos a Política de IA na SEGES, organizando os tópicos por capítulos.

GLOSSÁRIO

Modelo de Linguagem: Um modelo de linguagem é um sistema de inteligência artificial capaz de reconhecer, compreender ou gerar texto em linguagem humana baseado em exemplos de treinamento.

Modelos de Linguagem Open Source: modelos de linguagem cujo código-fonte é disponibilizado de forma aberta ao público, de modo que qualquer pessoa pode acessar, utilizar, modificar e distribuir esse código-fonte de forma livre.

Inteligência Artificial Generativa: é um tipo de IA que pode criar conteúdo e ideias, incluindo conversas, histórias, imagens, vídeos e músicas.

Opt-out: Refere-se à opção oferecida aos usuários para recusarem ou desativarem sua participação em determinadas atividades automatizadas, como o armazenamento de dados pessoais ou o uso de suas ações para fins de aprendizado de máquina. Esse mecanismo garante maior controle sobre a privacidade e segurança das informações, permitindo que o usuário decida se deseja ou não compartilhar seus dados com sistemas de IA.

Plataformas Públicas de IA generativa: Soluções de IA generativa, fornecidas por terceiros. Não confundir com gratuitas. Atualmente, são exemplos desses tipos de solução:

- ChatGPT, Copilot, Llama, Claude, Perplexity, Grok, Gemini, Maritaca AI e outras, que fornecem respostas em linguagem natural sobre basicamente qualquer tipo de assunto a partir de um prompt inicial;
- GitHub Copilot e Amazon CodeWhisperer, que produzem códigos de programação a partir de prompts;
- DALL-E, Midjourney e Stable Diffusion, que geram imagens a partir de prompts ou mesmo de outras imagens.

Prompt: Comando em linguagem corrente, normalmente em texto, apresentado a um modelo de linguagem de IA para gerar uma resposta ou realizar uma tarefa específica. A qualidade e precisão da resposta dependem diretamente da formulação do prompt.

Alucinação: termo usado na IA generativa para descrever respostas fictícias, confiantes e convincentes, que podem ser erroneamente aceitas devido a algum viés e assim escapar a uma revisão superficial.

Viés algorítmico: refere-se à presença de distorções sistemáticas nos resultados gerados por sistemas de inteligência artificial, resultantes de dados enviesados, falhas no design do modelo ou decisões arbitrárias no treinamento.

Esse viés pode se manifestar de diversas formas, incluindo discriminação racial, de gênero, socioeconômica, entre outras.

CAPÍTULO I – DOS OBJETIVOS

A Política de IA visa promover o uso ético, transparente e eficiente de tecnologias baseadas em IA, com foco na inovação responsável, proteção de direitos fundamentais e melhoria dos serviços públicos. Este manual detalha ações práticas para alcançar esses objetivos, como a criação de planos de adoção, capacitação institucional e monitoramento contínuo dos sistemas de IA.

CAPÍTULO II – DO CAMPO DE APLICAÇÃO

Aplica-se a todas as unidades da SEGES, abrangendo servidores, estagiários, fornecedores e parceiros envolvidos em projetos com IA. Recomenda-se que cada unidade identifique seus processos que utilizam ou pretendem utilizar IA e registre essas iniciativas em um inventário institucional de soluções automatizadas.

Da mesma forma, é recomendável que cada solução de IA, particularmente de IA generativa de uso público, seja avaliada no tocante aos benefícios e aos riscos que sua utilização possa trazer, antes de sua eventual adoção por uma área específica da Secretaria.

CAPÍTULO III – DOS PRINCÍPIOS ORIENTADORES

Importante recordar, também, os valores pelos quais a Administração pública - e a SEGES em particular – deve se pautar. Quais sejam a transparência, a publicidade, a moralidade, a economicidade, a impessoalidade e a não-discriminação, além da proteção aos dados pessoais de servidores, prestadores de serviços e munícipes.

Em decorrência dos princípios norteadores acima, recomendam-se alguns cuidados essenciais, por parte dos colaboradores da SEGES, na verificação de resultados gerados por ferramentas de IA generativa.

Entre esses cuidados, destacamos:

- Verificar se os resultados produzidos por ferramentas de IA generativa revelam algum tipo de viés discriminatório contra grupos sociais específicos, por motivos de etnia, gênero, orientação sexual, opinião política, credo religioso, filiação político-partidária ou sindical, condição sócio-econômica, condição de saúde, presença de algum tipo de deficiência ou qualquer outro fator que possa dar margem a preconceito e discriminação, violando assim princípios de não-discriminação, moralidade e impessoalidade da Administração;
- Caso se note que os resultados da IA generativa incorrem, de algum modo, em qualquer tipo de preconceito ou discriminação, esses resultados deverão ser ou corrigidos para a eliminação do problema, ou descartados;
- Verificar se os resultados gerados fazem sentido dentro do contexto e da questão formulada;
- Verificar se a resposta dada à questão e ou demanda proposta é internamente coerente, isto é, se as partes da resposta fazem sentido quando tomadas juntas;
- Verificar se essa mesma resposta coincide com verificações feitas de forma independente, via fontes confiáveis.
- Na mesma linha do item anterior, ter sempre cuidado especial com eventuais “alucinações” de IA.

Note-se que os cuidados propostos acima reforçam a necessidade de revisão humana dos resultados gerados pela IA, especialmente quando se utiliza a IA como suporte em decisões estratégicas e análise de situações e cenários complexos, onde os fatores a seguir exercem papel proeminente:

- Considerações de ordem prática
- Conhecimentos advindos de experiência
- Conhecimento técnico e jurídico especializado
- Nuances e sutilezas de ordem pessoal, cultural, política, social, etc.

Os fatores acima podem não ser sempre compreendidos pela IA generativa em seu estágio atual de desenvolvimento, de modo que as situações onde tais fatores estão envolvidos poderão ser melhor interpretadas e resolvidas com a expertise humana, com a IA tendo papel de suporte e assistência.

Reforçamos, ainda mais uma vez que, especialmente para temas técnicos ou legais que demandam expertise humana, deve-se confirmar a veracidade dos resultados gerados pela IA por meios independentes e confiáveis.

Por fim, recomenda-se a não utilização de ferramentas de IA generativa pelo servidor em casos onde este acredite não possuir conhecimento técnico necessário para a avaliação dos resultados gerados pela IA.

CAPÍTULO IV – DAS BOAS PRÁTICAS DE SEGURANÇA NO USO DA IA GENERATIVA

Reforçando a necessidade de se prestar atenção à segurança dos usuários, dados e sistemas de informação institucionais, tecemos as seguintes recomendações:

- Evitar criar contas em ferramentas públicas de IA generativa utilizando e-mails, telefones e demais credenciais institucionais;
- Dados sensíveis só podem ser utilizadas por IAs que tenham recursos para utilização corporativa e que mantenham o compromisso de não-vazamento de dados e não-aprendizado público, e deverão ser utilizados estritamente dentro dos limites legais, para a consecução das atribuições da Secretaria;
- Lembrar sempre que informações sensíveis, protegidas ou sigilosas da SEGES ou da Administração municipal como um todo, se fornecidas indevidamente a ferramentas não-aprovadas de IA generativa, ou que não tenham compromisso de não-vazamento e não-aprendizado público, poderão eventualmente ser reveladas para outros usuários, configurando vazamento de dados, com consequências legais, reputacionais e financeiras potencialmente sérias para a Administração municipal.
- Da mesma forma, evitar inserir dados em ferramentas de IA generativa não-aprovadas na SEGES que possam permitir que tais ferramentas realizem inferências sobre informações confidenciais ou classificadas, considerando o histórico de interações do usuário;
- Ter sempre cuidado com links fornecidos por ferramentas públicas de IA generativa, especialmente aquelas ferramentas ainda não aprovadas para uso interno na Secretaria. Esses links podem ser maliciosos. Preferir sempre acessar fontes confiáveis e verificadas antes de clicar em qualquer link;

- Em hipótese alguma compartilhar informações pessoais de servidores, munícipes ou terceiros em ferramentas de IA generativa públicas, exceto quando expressamente autorizado pela política de proteção de dados da SEGES ou da Administração municipal. Dados publicamente disponíveis na internet podem ser utilizados conforme a legislação aplicável.
- Por segurança, também evitar o uso de logins e senhas institucionais ao registrar contas em ferramentas públicas de IA generativa. Utilizar, quando possível, autenticação de múltiplos fatores (MFA) ou outras camadas de segurança disponíveis, para aumentar a segurança de acesso;
- Em consonância com as observações de Capítulos anteriores, em especial do Capítulo III, revisar cuidadosamente os conteúdos gerados pela IA generativa antes de compartilhá-los interna ou externamente, garantindo assim a conformidade com as diretrizes institucionais e evitando a disseminação de informações imprecisas, errôneas ou sensíveis;
- Informações sensíveis geradas por IA não devem ser armazenadas inadequadamente, devendo haver locais seguros para seu armazenamento;
- Estar sempre atualizado sobre as políticas de IA generativa da SEGES, que poderão sofrer alterações a qualquer tempo, uma vez que as ferramentas e possibilidades de uso da IA generativa estão em rápida evolução.

CAPÍTULO V – DAS DIRETRIZES GERAIS

- 1. Os projetos de IA deverão ser avaliados pela área técnica competente, em conjunto com a área de TIC da Secretaria;
- 2. Poderá ser criado - a qualquer tempo e conforme a necessidade e conveniência da Secretaria - um Plano de Adoção e Monitoramento de Ferramentas de IA em SEGES, que poderá conter métricas específicas, como o número de sistemas e/ou fluxos de trabalho com utilização de ferramentas de IA, as capacitações realizadas por servidores da SEGES, dentre outros indicadores que se façam relevantes. O referido Plano, se implementado, poderá ser atualizado a qualquer momento.
- 3. Poderá ser realizada, pela área técnica usuária e pela COTIC, uma classificação de risco para novas soluções de IA que vierem a ser testadas e utilizadas em SEGES. A classificação de risco seguirá as orientações constantes do Manual de Gestão de Riscos.
- 4. Os contratos para soluções de IA que vierem a ser celebrados pela Secretaria deverão seguir o disposto no Capítulo V deste documento.
- 5. Sempre que possível, priorizar soluções de IA com código aberto e com documentação pública.

CAPÍTULO VI – DOS REQUISITOS TÉCNICOS MÍNIMOS

Toda solução de IA que vier a ser contratada em SEGES junto ao mercado, ou mediante acordo de cooperação técnica com outro órgão ou instituição, pública ou privada, deverá conter documentação técnica que descreva, ao menos:

- O modelo de linguagem de IA;
- As capacidades esperadas da ferramenta – o que a ferramenta pode e não pode fazer, no momento da contratação.
- Forma de tratamento de dados pessoais, especialmente dados pessoais sensíveis, observados os segredos industrial e comercial;
- Se a ferramenta possui políticas e mecanismos de não-utilização e não-aprendizado de dados considerados sensíveis pela SEGES, sejam esses dados pessoais ou não. Um exemplo é a possibilidade de oferecimento de opt-out para os usuários.

Opcionalmente, a documentação técnica da solução de IA poderá conter:

- Quais bases de dados foram utilizadas no treinamento do modelo de IA, respeitados eventuais segredos industriais e comerciais;
- Dados ou estimativas sobre consumo energético da ferramenta de IA, caso estejam disponíveis.

A interoperabilidade entre a solução de IA contratada deverá ser testada com os sistemas da PMSP com os quais a solução interaja, e a explicabilidade deve ser garantida por meio de interfaces que permitam ao usuário compreender as decisões automatizadas.

CAPÍTULO VII – DOS DIREITOS DO CIDADÃO E DO USUÁRIO

Os sistemas de IA devem conter mecanismos que permitam ao cidadão:

- Saber quando está interagindo com IA, posto que a divulgação deste fato atende aos princípios da publicidade e da transparência da Administração pública;
- Solicitar revisão humana de decisões automatizadas;
- Corrigir informações incorretas;
- Ter seus dados protegidos conforme a LGPD;
- Optar por outros canais de atendimento disponíveis que não façam uso de IA.

CAPÍTULO VIII – DAS RESPONSABILIDADES E MECANISMOS DE CONTROLE

Recomenda-se que cada unidade da SEGES designe um ou mais pontos focais para assuntos relacionados à IA, de forma que as responsabilidades de controle fiquem assim distribuídas:

- A COTIC manterá registros dos sistemas em operação e garantirá a conformidade técnica.
- A AJ revisará os contratos e pareceres quanto aos aspectos legais e jurídicos.
- A COGEP supervisionará o uso de IA em processos de RH.
- Poderá, a critério da SEGES, haver algum tipo de auditoria do uso das soluções contratadas, e esta auditoria, se implementada, poderá ser realizada pelos meios apropriados e disponíveis na ocasião.

CAPÍTULO IX – DA CAPACITAÇÃO E CULTURA INSTITUCIONAL

As capacitações que forem realizadas por servidores da SEGES que trabalhem com IA incluirão, tanto quanto possível, trilhas formativas sobre:

- Fundamentos técnicos;
- Como aprimorar os resultados gerados pela IA;
- Ética em IA;
- Boas práticas e recomendações de segurança no uso de IA;
- Legislação aplicável à IA.

Recomenda-se a realização de oficinas práticas, produção de cartilhas e seminários anuais.

Parcerias com universidades e startups poderão ser formalizadas por meio de termos de cooperação técnica.

CAPÍTULO X – DA PARTICIPAÇÃO E CONTROLE SOCIAL

Poderá ser produzido, conforme a necessidade da SEGES, um Relatório de Transparência e Impacto da IA, em momento que a SEGES julgar oportuno, e a publicação de tal relatório será, preferencialmente, no portal institucional da SEGES, contendo dados sobre sistemas em uso, riscos identificados e medidas de contenção de risco adotadas.

Irregularidades no uso de ferramentas de IA serão averiguadas pela área técnica responsável pela ferramenta e, se necessário, encaminhadas ao gabinete da Secretaria para eventuais medidas posteriores.

CAPÍTULO XI – DA UTILIZAÇÃO PRÁTICA DA IA, DA CAPACITAÇÃO E DA NECESSIDADE DE ATUALIZAÇÃO

Atualmente, existem muitas possibilidades de utilização para a IA generativa no setor público, tais como tradução e revisão de documentos, resumo de documentos extensos, apoio à redação e confecção de documentos, classificação e categorização de dados, processamento de dados em massa, consultas a documentos e normativos, geração de arquivos de mídia, transcrições de áudios e vídeos, automação de processos internos e mesmo, em algumas circunstâncias, do atendimento ao cidadão, e até mesmo geração de código, dentre muitas outras possibilidades.

Diante de possibilidades cada vez maiores de utilização de IA generativa dentro dos fluxos de trabalho da SEGES, recomendamos que os servidores, na medida do possível e das necessidades específicas de cada um, engajem-se proativamente em capacitações sobre IA generativa, especialmente aquelas fornecidas por entidades integrantes da Administração pública, como as escolas de governo, ou por entidades externas que tenham algum tipo de parceria com a Administração municipal para fins de capacitação de servidores.

Recomendamos ainda que os servidores que fazem uso de ferramentas de IA generativa, quer de forma frequente ou ocasional, mantenham-se, tanto quanto possível, atualizados sobre o tema.

CAPÍTULO XII – DA UTILIZAÇÃO DA IA GENERATIVA PARA PRODUÇÃO DE CÓDIGO

Como levantado no Capítulo anterior, existe a possibilidade, atualmente, do uso da IA generativa para elaboração de códigos em programação. Sobre este assunto, tecemos as seguintes recomendações:

Todo código de programação produzido por IA generativa deve ser revisado por especialistas de TI antes de qualquer implementação prática na instituição, para garantir a qualidade, segurança e conformidade com padrões técnicos e legais;

É fundamental monitorar continuamente o desempenho de códigos gerados por IA em relação à finalidade pretendida, pois os modelos de IA podem ser atualizados periodicamente e aprender com os dados fornecidos, o que pode impactar sua confiabilidade e precisão.

Usar o recurso opt-out (exclusão voluntária), quando couber, para reduzir a utilização dos prompts no treinamento e desenvolvimento de um sistema de IA.

CAPÍTULO XIII – DOS DIREITOS AUTORAIS

Recomendamos evitar, tanto quanto possível, o uso de conteúdo protegido por direitos autorais, tais como imagens, vídeos, textos, áudios, dentre outros, quando do uso de aplicações de IA generativa, em material institucional. Tal medida visa resguardar a Secretaria e seus servidores de possíveis consequências administrativas e legais.

Se houver dúvida quanto ao fato de que determinado material gerado por IA generativa possa estar protegido por direitos autorais, recomenda-se não utilizar tal material.

Em casos onde for preciso utilizar conteúdo protegido por direitos autorais, dar sempre preferência para fontes que sejam de notório domínio público, e sempre citar o(s) autor(es) e/ou fonte(s) do material, dando-lhes o crédito devido.

CAPÍTULO XIV – DAS DISPOSIÇÕES FINAIS E TRANSITÓRIAS

O presente manual funcionará como uma diretriz para assuntos de IA dentro da SEGES e que poderá ser revisado e atualizado a critério da área de TIC da Secretaria, sempre que esta julgar necessário.



**PREFEITURA DE
SÃO PAULO**