MANIFESTAÇÃO TÉCNICA

Nestes autos do Processo de Gestão Administrativa nº 001.2023.031833, vimos apresentar manifestação técnica, de caráter opinativo, acerca da utilização de ferramentas de inteligência artificial, a exemplo do ChatGPT, no exercício da atuação do Ministério Público e para o desempenho de suas funções essenciais.

Em princípio, quando tratamos de tecnologia e inovação e de sua aplicação em meio institucional e/ou corporativo, o recomendado é iniciar os estudos de utilização em um escopo hermético e controlado, implementando e avaliando seu impacto na rotina sem descuidar das questões atinentes à privacidade e segurança, até que se adquira a maturidade e a confiança necessárias para utilização em definitivo da tecnologia como suporte na tomada de decisões com base nos resultados obtidos através da interação com essas ferramentas.

No cenário ministerial, por sua vez, as ferramentas de inteligência artificial não substituem a atuação do membro no desenvolvimento de suas funções essenciais, mas constituem meio auxiliar que pode trazer maior celeridade em atividades rotineiras e operacionais da instituição.

Embora as ferramentas de Processamento de Linguagem Natural (PLN) não constituam necessariamente uma novidade, é notória sua recente popularização desde o lançamento do ChatGPT em novembro de 2022, pela startup americana OpenAI, desde então projetando a temática na pauta de discussões das organizações da iniciativa privada, bem como do setor público, quanto à sua utilização para melhorar processos de negócio, promover eficiência e reduzir custos. Nesse aspecto, apontem-se os princípios constitucionais administrativos previstos no art. 37 e economicidade, previsto no art. 70, ambos da Constituição Federal.

O ChatGPT é baseado em tecnologia de *machine learning* e foi projetado para melhorar a comunicação homem-máquina, podendo realizar várias tarefas como: geração e/ou predição de texto, responder questões, sumarizar documentos tornando-os mais concisos, traduzir documentos, classificar textos e realizar a análise de sentimento. É incontestável que representa uma ferramenta de grande potencial, porém é importante considerar suas limitações e potenciais riscos.

Especificamente sobre esses tipos de ferramenta, o principal risco imposto por eles consiste na utilização de **algoritmos mal treinados,** que podem trazer informações incorretas, instruções prejudiciais ou conteúdo tendencioso. Esse

tipo de problema é denominado "*desalinhamento(misalignment)*" e consiste em um maior risco em que a máquina realiza a tomada de decisão sem supervisão humana. Dessa maneira, no contexto ministerial, as respostas fornecidas por uma inteligência artificial necessariamente requerem um olhar humano especializado, sob o risco de comprometer a reputação da pessoa que as utiliza ou até mesmo da instituição que aquele representa caso o resultado da inteligência seja meramente transcrito em um documento sem qualquer avaliação em relação ao seu conteúdo.

Especificamente quanto aos modelos de linguagem de grande escala, *"large language models" (LLM),* há o risco de **disseminação de informação inverídica**, tendo em vista que a compreensão dos modelos é limitada e podem estar alimentados com informações desatualizadas, ambíguas ou parcialmente compreendidas. Uma vez que um dos conjuntos de dados utilizados no treinamento do ChatGPT foram informações obtidas diretamente da Internet, via de regra, nem sempre se terá a resposta correta para todas as questões propostas, muito embora possam parecer plausíveis. O fato de estar utilizando uma tecnologia de vanguarda não implica na confiabilidade das respostas, e o próprio ChatGPT faz menções a esse respeito em determinadas circunstâncias. Essa problemática é definida como "*alucinação*"do modelo.

Questões de cunho ético e legal também merecem destaque, tendo em vista que as respostas oferecidas pela ferramenta de IA podem estar em **desconformidade com fatores éticos, sociais, culturais, ambientais e de *compliance***, o que também representa um risco. Vale ressaltar que o fluxo conversacional entre a IA e o usuário é estabelecido conforme os *inputs* fornecidos e, na simulação de uma conversa humana, o modelo de aprendizado de máquina pode equivocadamente persuadir a pessoa com dados incoerentes.

No entanto, como forma de mitigar esses riscos algumas medidas podem ser adotadas. O advento dessas tecnologias em meio institucional e corporativo faz despertar para a responsabilidade das auditorias e áreas de gestão de riscos e para a necessidade de definição e implementação de processos claros, bem como de seu monitoramento, e a elaboração de normativos e regulamentações para implementação e uso considerando políticas de segurança e privacidade, sobretudo as que se comunicam com a preservação de dados pessoais.

Um maior envolvimento interinstitucional e com centros de pesquisa em inovação e tecnologia pode ser útil no sentido de conjugar esforços e compartilhar experiências. Aliado a isso, a comunicação institucional deve ser estimulada e ações educativas sobre condutas seguras, assim como informativos sobre os riscos de compartilhamento de informações pessoais e sensíveis podem ser trabalhados em conjunto.

A despeito dos riscos apresentados, é importante, por fim, ressaltar que a adoção de ferramentas de PLN em âmbito ministerial pode trazer benefícios, tais como os indicados na Nota Técnica n° 1/2023 que dispõe sobre o Uso Seguro de Ferramentas de Processamento de Linguagem Natural no Ministério Público Brasileiro. São eles: redução do tempo de resposta, melhoria na qualidade das análises e decisões, otimização dos recursos humanos, redução de custos, aumento

da capacidade de atendimento e priorização de temas importantes, inovação e modernização e incremento de eficiência para a proteção de dados pessoais, representando um passo significativo para os avanços da inteligência artificial e nos meios em que esta tecnologia pode ser utilizada.

Sob o aspecto da Lei Geral de Proteção de Dados, a utilização de LLMs já treinados não apresenta risco maior do que os modelos convencionais de aprendizado de máquina já utilizados no mercado, pois se trata de um modelo fundacional (*foundation model*), com aprendizado auto supervisionado ou semi-autosupervisionado. Nessas situações, a retenção de dados dos usuários, durante a sua utilização não se mostra tão importante ou gera uma grande melhoria na qualidade geral do LLM do que a obtenção de novos conjuntos de dados oriundos de outras fontes. Ao contrário, existe grande valia em conjuntos de dados categorizados para treinamento de machine learning para finalidades específicas, a exemplo de categorização. Esses processos podem podem ser bastante aprimorados com a resposta positiva ou negativa dos usuários. Embora exista a possibilidade de "*finetuning*" de um LLM, para casos específicos, essa necessidade atende mais às expectativas do usuário do que o aprimoramento do produto pelas empresas.

Outro aspecto a considerar é a publicidade dos processos e procedimentos do Ministério Público. Nesse aspecto aponte-se inclusive que o próprio Estatuto da Advocacia autoriza o exame de qualquer procedimento ou processo sem necessidade de apresentação de justificativa, ou procuração de cliente, exceto se sigiloso.

Nesse sentido, a própria OAB não proibiu advogados de utilizarem essas ferramentas e tem incentivado o seu uso, como demonstra o seguinte link: https://jornaldaadvocacia.oabsp.org.br/noticias/comissao-de-ia-da-oab-sppromove-seminario-inedito-sobre-chatgpt-no-ambiente-juridico/. O próprio representante, até onde se tem conhecimento, não buscou a aplicação de medida proibitiva equivalente perante o seu órgão de representação.

João Pessoa, 09 de maio de 2023.

(assinado eletronicamente)
Alberto Vinicius Cartaxo da Cunha
Promotor de Justiça
Encarregado pelo Tratamento de Dados Pessoais - MPPB

(assinado eletronicamente)
Vivianne de Queiroz Leal
Analista Ministerial
Diretora de Tecnologia da Informação