

DIRETRIZES PARA UTILIZAÇÃO INSTITUCIONAL DAS FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL

DIRETRIZES PARA UTILIZAÇÃO INSTITUCIONAL DAS FERRAMENTAS DE INTELIGÊNCIA ARTIFICIAL

ITAJUBÁ - MG

JULHO/2024

DIRETORA GERAL

Cristiane Resende

COORDENADORA ACADÊMICA

Talyta Resende de Oliveira

COORDENADORA ADMINISTRATIVO-FINANCEIRA

Karen Bianca Dias Ribeiro

**COORDENADORA DE PESQUISA, EXTENSÃO, INTERNACIONALIZAÇÃO E
INOVAÇÃO**

Renata de Castro Matias

PROCURADORA INSTITUCIONAL

Talyta Resende de Oliveira

NÚCLEO DE INOVAÇÃO ACADÊMICA

Gerson Hiroshi Yoshinari Júnior

NÚCLEO DE PESQUISA

Luciano Magalhães Vitorino

Versão atualizada em julho de 2024

1. Introdução

A implementação de diretrizes para o uso de ferramentas de Inteligência Artificial (IA) na Faculdade de Medicina de Itajubá (FMIT-AFYA) é crucial para assegurar um uso equilibrado, ético e inovador dessas tecnologias. Em cursos da área da saúde, como medicina e enfermagem a IA tem o potencial de revolucionar o ensino, a pesquisa e a extensão universitária. No entanto, é essencial que seu uso seja regulado para maximizar benefícios e minimizar riscos. A atenção do Núcleo de Pesquisa (NUP-FMIT) e do Núcleo de Inovação Acadêmica (NIA-FMIT) estabelecer esses critérios demonstra um compromisso com a inovação responsável. Diretrizes bem definidas garantem que a IA seja utilizada para melhorar a eficiência administrativa, personalizar o ensino e facilitar a pesquisa acadêmica, ao mesmo tempo que protegem a privacidade e a integridade acadêmica, prevenindo o uso inadequado. Isso promove um ambiente acadêmico de alta qualidade, preparado para os desafios do futuro.

1.1 Objetivo

O objetivo desta diretriz é estabelecer orientações claras para o uso ético e eficaz de ferramentas de Inteligência Artificial (IA), com ênfase em modelos de Processamento de Linguagem Natural (PLN), na FMIT-AFYA. A IA tem o potencial de transformar a educação, pesquisa e administração, mas também apresenta desafios éticos e legais que devem ser abordados de forma proativa.

1.2 Contextualização Básica

1.2.1 O que são Ferramentas de Inteligência Artificial?

Ferramentas de IA são sistemas computacionais que imitam habilidades humanas como aprendizado, raciocínio e resolução de problemas. Modelos de PLN, um subconjunto da IA, são projetados para entender, interpretar e gerar texto de forma semelhante à humana.

1.2.2 Aplicabilidade na Educação.

A IA pode ser integrada de forma significativa no design educacional para apoiar o letramento digital e o desenvolvimento do pensamento crítico. Por exemplo, modelos de *chatbot* podem ser usados em reuniões de tutoria para *brainstorming* e elaboração de ensaios iniciais. Além disso, os alunos podem ganhar insights e experiência em funcionalidades de IA enquanto abordam criticamente a saída de IA ("output").

1.3 Potenciais Benefícios

Melhoria da eficiência administrativa

Personalização do ensino

Facilitação da pesquisa acadêmica

Apoio à tomada de decisão

1.4 Potenciais Riscos

Questões de privacidade e consentimento

Riscos de plágio e fraude acadêmica

Desafios na avaliação do desempenho da aprendizagem dos alunos

Certificação da veracidade da informação

A Maastricht University destaca que o design educacional em Aprendizagem Baseada em Problemas (PBL) pode integrar de forma significativa ferramentas de IA para apoiar o letramento digital e o desenvolvimento do pensamento crítico. Além disso, essas ferramentas desafiam a forma como as universidades podem medir o sucesso dos alunos em alcançar os resultados de aprendizagem pretendidos, especialmente para avaliações escritas em um ambiente não controlado.

2. Escopo

O escopo desta normatização abrange a utilização de ferramentas de IA dentro da FMIT-AFYA. Embora as diretrizes incluam ênfase em modelos de PLN, elas não se limitam a esses modelos. O escopo também abrange todos os modelos generativos, incluindo geradores de imagem, vídeo e som. A normatização é aplicável a todos os membros da comunidade acadêmica, incluindo estudantes, professores e pessoal técnico-administrativo.

2.1 O documento visa estabelecer diretrizes claras para:

Uso Ético: Garantir que o uso de IA esteja alinhado com princípios éticos, incluindo questões de consentimento informado e privacidade.

Integridade Acadêmica: Prevenir a fraude acadêmica, assegurando que as ferramentas de IA não sejam usadas para completar avaliações ou tarefas de forma desonesta.

Responsabilidade e Consequências: Definir as responsabilidades dos usuários e as consequências para o uso inadequado das ferramentas.

Ao focar nesses aspectos, a normatização busca evitar o mau uso das ferramentas de IA e garantir que elas sejam usadas de forma a maximizar os benefícios enquanto minimiza os riscos associados.

3. Princípios Éticos e Legais

Este documento estabelece normas para o uso ético e legal de ferramentas de IA e modelos de PLN na FMIT-AFYA. As diretrizes são fundamentadas em princípios éticos e legais que visam garantir a integridade, privacidade e inclusão. O

documento "What's in the Chatterbox? Large Language Models, Why They Matter, and What We Should Do About Them" da Universidade de Michigan destaca várias implicações éticas e sociais do uso de modelos de linguagem grandes (LLMs). Entre as preocupações estão a reprodução de preconceitos históricos, o uso excessivo de recursos e a concentração do desenvolvimento de LLMs em um pequeno número de grandes empresas de tecnologia. Estas questões reforçam a necessidade de diretrizes éticas rigorosas para o uso de IA e PLN em ambientes acadêmicos.

3.1 Uso Ético

O uso ético de IA e PLN deve ser alinhado com princípios de consentimento informado, privacidade e não-discriminação. As ferramentas não devem ser usadas para fins de desinformação, plágio ou qualquer forma de fraude acadêmica.

O Uso ético, baseado no documento "Recomendação sobre a Ética da Inteligência Artificial", publicado pela Unesco em 2021, traz os seguintes princípios norteadores:

3.1.2 Privacidade e Consentimento

Garantir a proteção da privacidade dos dados dos usuários (alunos, professores, funcionários).

Obter consentimento informado para coleta e uso de dados, especialmente dados sensíveis.

3.1.3 Transparência

Ser claro sobre como e por que os dados são coletados e utilizados. Fornecer informações sobre como as ferramentas de IA estão sendo implementadas e utilizadas.

3.1.4 Responsabilidade

Identificar partes responsáveis pela implementação e manutenção das ferramentas de IA. Estabelecer procedimentos claros para a responsabilização em casos de má utilização ou falhas, incluindo responsabilização legal.

3.1.5 Integridade e Confidencialidade

Proteger a integridade dos dados e garantir a confidencialidade das informações. Desenvolver e implementar políticas claras para prevenir a manipulação e o acesso não autorizado aos dados.

3.1.6 Equidade e Inclusão

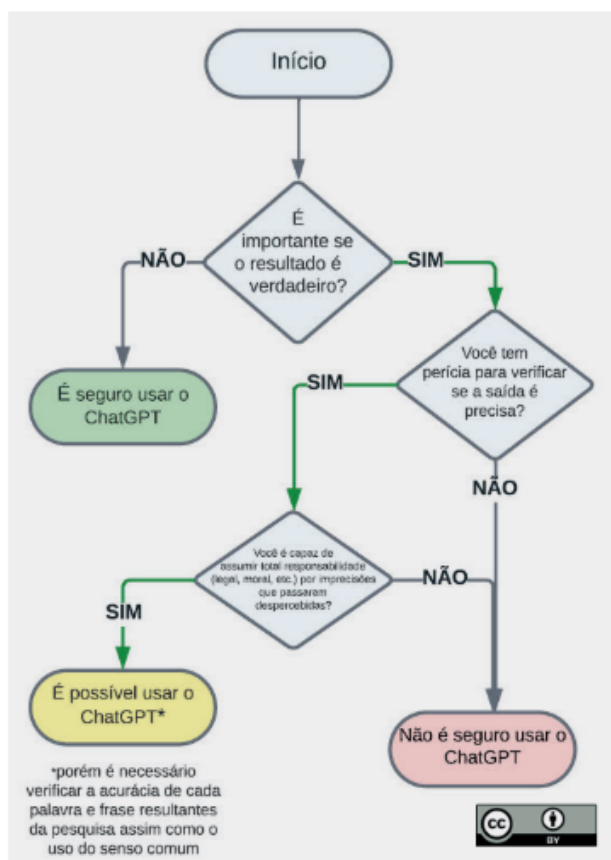
Garantir que as ferramentas de IA sejam inclusivas e acessíveis a todos os usuários.

Trabalhar para minimizar e mitigar possíveis vieses em ferramentas de IA.

3.1.7 É seguro utilizar a IA na educação?

A Figura 1 abaixo apresenta um fluxograma que auxilia na determinação das condições seguras para o uso de uma IA, como por exemplo o ChatGPT. Este fluxograma orienta os usuários a considerarem vários fatores antes de utilizar a ferramenta, incluindo a natureza da informação solicitada, a sensibilidade dos dados e a finalidade do uso. É crucial garantir que as consultas sejam apropriadas e que o ChatGPT não seja usado para decisões críticas sem supervisão humana. Além disso, é importante estar ciente das limitações do modelo, que não pode verificar a veracidade das informações fornecidas ou avaliar sua própria confiabilidade. Portanto, o uso seguro do ChatGPT deve envolver a avaliação crítica das respostas e a consideração dos contextos éticos e de privacidade.

Figura 1: Quando é seguro utilizar o ChatGPT?⁵



Fluxograma realizado por Aleksandr Tiulkanov, advogado para política de dados e IA, janeiro de 2023.

Disponível em: ChatGPT e Inteligência Artificial na Educação Superior: Guia de Início Rápido. UNESCO; 2023.

3.1.8 Educação e Conscientização.

Promover a educação e a conscientização sobre o uso ético da IA entre alunos, professores e funcionários. Fornecer treinamento e recursos educacionais sobre o uso responsável de tecnologias de IA.

3.1.9 Inovação e Desenvolvimento Sustentável

Encorajar o uso de IA para inovar e melhorar os processos educativos. Alinhar o uso de IA com os objetivos de desenvolvimento sustentável e ético.

3.1.10 Avaliação e Melhoria Contínua

Monitorar regularmente o impacto das ferramentas de IA e fazer ajustes conforme necessário.

Estabelecer mecanismos de feedback para identificar áreas de melhoria e ajustar práticas conforme necessário.

3.1.11 Colaboração Interdisciplinar

Promover a colaboração entre diferentes disciplinas para garantir uma abordagem holística no desenvolvimento e implementação de IA. Envolver especialistas em ética, direito, tecnologia da informação, medicina e outras áreas relevantes.

3.1.12 Bem-Estar dos Usuários

Priorizar o bem-estar dos usuários ao implementar ferramentas de IA. Considerar os impactos psicológicos e emocionais do uso de IA na comunidade educacional.

4. Diretrizes para Utilização de IA e PLN

O objetivo destas diretrizes é assegurar que os usuários assumam total responsabilidade sobre as informações obtidas através de ferramentas de IA e PLN. Um estudo da Universidade de Oxford destaca que, enquanto os usuários não podem assumir todo o crédito por resultados positivos gerados por LLMs, é apropriado responsabilizá-los por usos prejudiciais, como a geração de desinformação ou a falta de cuidado na verificação da precisão do texto gerado.

4.1 Integridade Acadêmica

A IA é de suma importância. O uso inadequado de IA e PLN para completar avaliações ou tarefas acadêmicas de forma desonesta é estritamente proibido.

4.2 Responsabilidade do Usuário na Checagem de Informações

É imperativo que os usuários verifiquem as informações obtidas através das ferramentas de IA e PLN junto a fontes confiáveis. Além disso, essas fontes devem ser devidamente citadas em qualquer documento ou pesquisa onde as informações sejam utilizadas. Também **é obrigatória a expressa comunicação de que ferramentas de IA foram utilizadas na elaboração de material acadêmico**. Como exemplo de comunicação desse tipo de conteúdo, trazemos o texto que compõem os materiais publicados pelo portal de informações médicas MEDSCAPE:

“Este artigo foi traduzido da edição francesa do Medscape usando várias ferramentas editoriais, incluindo IA, como parte do processo. Editores humanos revisaram este conteúdo antes da publicação.”

4.3 Uso por Estudantes em Situações Avaliativas e outras atividades educacionais

O uso de ferramentas de IA e PLN por estudantes é proibido em situações avaliativas onde suas habilidades de retenção de conhecimento estão sendo exclusivamente avaliadas.

4.4 Ferramentas de Deepfake

Deepfake é uma tecnologia que utiliza IA para criar representações realistas, mas falsas, de pessoas em fotos, vídeos ou áudio. Esta tecnologia pode ser usada para criar conteúdo enganoso e é terminantemente proibida na instituição.

4.5 Responsabilidade na Curadoria de Informações

Os usuários que são colaboradores da instituição têm a responsabilidade de curadoria das informações no que diz respeito às políticas e diretrizes específicas da instituição. Estas políticas e diretrizes são amplamente divulgadas por meio de treinamentos obrigatórios. Ao seguir estas diretrizes, os usuários contribuem para um ambiente acadêmico mais ético e responsável, alinhado com os princípios e valores da instituição e da sociedade em geral.

5. Responsabilidades e Consequências

Esta seção delinea as responsabilidades e as possíveis consequências para o uso inadequado de ferramentas de IA e PLN dentro da FMIT-AFYA.

5.1 Investigação e Sindicância

Casos de uso inadequado das ferramentas identificados por quaisquer membros da instituição, de forma pública ou anônima via ouvidoria, serão investigados pela coordenação de curso através de sindicância interna. Após devida apuração, a coordenação poderá oferecer denúncia ao colegiado da faculdade, que decidirá as medidas cabíveis à sindicância.

5.2 Infrações de Ética Médica

Possíveis infrações de ética médica identificadas durante a sindicância serão encaminhadas para o conselho de medicina para avaliação e ação apropriada.

5.3 Ilegalidades

Na constatação de possíveis ilegalidades, as informações serão encaminhadas aos órgãos competentes para investigação e ação legal.

5.4 Cenário em Mudança

Reconhecemos que o cenário de desenvolvimento e utilização de ferramentas de IA é altamente mutável. A instituição estará atenta a essas mudanças e convida toda a comunidade acadêmica a debater de forma construtiva sobre o assunto. O objetivo é ajudar a construir diretrizes que primam pela utilização responsável de tecnologias que sejam benéficas a todos. Ao estabelecer um mecanismo claro de responsabilidades e consequências, esta seção visa garantir que o uso de IA e PLN esteja alinhado com os princípios éticos e legais, tanto da instituição quanto da sociedade em geral.

6. Disposições finais

É importante salientar que este documento será revisado e atualizado conforme novas legislações específicas sobre o uso de AI e PLN entrarem em vigor. A instituição se compromete a manter as diretrizes alinhadas com as normas legais vigentes, garantindo assim a conformidade e a responsabilidade no uso dessas tecnologias.

6.1 Conformidade com a LGPD

É crucial que todos os usuários estejam cientes da Lei Geral de Proteção de Dados (LGPD)

no Brasil ao utilizar ferramentas de AI e PLN. A LGPD estabelece diretrizes rigorosas para a coleta, armazenamento e uso de dados sensíveis. O não cumprimento desta legislação pode resultar em sérias consequências legais. Portanto, é imperativo que os usuários sigam as normas estabelecidas pela LGPD ao interagir com essas tecnologias.

6.2 Proteção de Informações Pessoais

A proteção de informações pessoais é de suma importância no uso de ferramentas de IA e PLN. Os usuários são fortemente aconselhados a evitar o uso de dados pessoais, tais como nomes, endereços e informações de saúde, nas ferramentas de IA e PLN.

7. Estudo de Engenharia de Prompt

Para maximizar o desempenho e a eficácia das ferramentas de IA e PLN sugerimos aos usuários que estudem técnicas de engenharia de prompt. Essas técnicas ajudam a formular perguntas e comandos de forma que as ferramentas possam fornecer respostas mais precisas e úteis. Existem diversos cursos

disponíveis na internet que oferecem treinamento em engenharia de prompt, e o investimento nesse conhecimento pode ser altamente benéfico.

Referência Bibliográfica

1. Universidade de Michigan. What's in the Chatterbox? Large Language Models, Why They Matter, and What We Should Do About Them. Abril de 2022.
2. Yoshinari Júnior GH, Vitorino LM. How may ChatGPT impact medical teaching? Arq Bras Med. 2023;10.1590/1806-9282.20230282. Disponível em: <https://doi.org/10.1590/1806-9282.20230282>.
3. Universidade de Oxford. Tackling the Ethical Dilemma of Responsibility in Large Language Models. Maio de 2023. Disponível em: <https://www.ox.ac.uk/news/2023-05-05-tackling-ethical-dilemma-responsibility-large-language-models>. Acesso em: 30 de agosto de 2023.
4. NATURE MACHINE INTELLIGENCE. Generative AI entails a credit-blame asymmetry. Disponível em: <https://www.nature.com/articles/s42256-023-00653-1>. Acesso em: 30 de agosto de 2023.
5. Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO). Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO; 2022. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000381137_por. Acesso em: 26 de setembro de 2023.
6. Sabzalieva E, Valentini A. ChatGPT e Inteligência Artificial na Educação Superior: Guia de Início Rápido. UNESCO; 2023. Documento ED/HE/IESALC/IP/2023/12. Disponível em: https://unesdoc.unesco.org/ark:/48223/pf0000385146_por.locale=en.

Glossário

Inteligência Artificial (IA): É a simulação de processos de inteligência humana por sistemas de computador. Inclui aprendizado, raciocínio e auto-correção.

Aprendizado de Máquina (Machine Learning): Subcampo da IA que se concentra no desenvolvimento de algoritmos que permitem aos sistemas aprender e melhorar a partir de dados, em vez de serem explicitamente programados.

Aprendizado Profundo (Deep Learning): Uma subcategoria do aprendizado de máquina que utiliza redes neurais profundas para modelar e resolver tarefas complexas, como reconhecimento de voz, visão computacional e processamento de linguagem natural.

Rede Neural (Neural Network): Um modelo computacional inspirado no funcionamento do cérebro humano, composto por neurônios artificiais interconectados. É usado para tarefas de aprendizado profundo.

Metodologia Transformer: Uma arquitetura de rede neural que se destacou no processamento de linguagem natural e em várias outras tarefas. Ela se baseia em atenção (attention) e é amplamente utilizada em modelos de linguagem grandes (LLMs).

Processamento de Linguagem Natural (PLN): Subcampo da IA que foca na interação entre computadores e linguagem humana. É a tecnologia por trás de assistentes virtuais, tradutores automáticos e outras aplicações que requerem compreensão da linguagem.

Modelos de Linguagem Grandes (LLMs): São modelos de IA especializados em tarefas de linguagem. Eles são treinados em grandes conjuntos de dados e são capazes de gerar texto que pode ser indistinguível do que um humano escreveria.

Deepfake: Tecnologia que utiliza IA para criar representações realistas, mas falsas, de pessoas em fotos, vídeos ou áudio.

Sindicância Interna: Processo de investigação conduzido pela coordenação de curso para apurar casos de uso inadequado das ferramentas de IA.

Ética Médica: Conjunto de princípios éticos que governam a conduta dos profissionais de medicina.

LGPD (Lei Geral de Proteção de Dados): Legislação brasileira que regula a coleta, armazenamento e uso de dados pessoais, estabelecendo diretrizes rigorosas para a proteção dessas informações.

Engenharia de Prompt: Técnica que ajuda a formular perguntas e comandos de forma que as ferramentas de IA possam fornecer respostas mais precisas e úteis.



FMIT | Afya