

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE GOIÁS
ESCOLA POLITÉCNICA E DE ARTES
BACHARELADO EM ENGENHARIA DA COMPUTAÇÃO



GUILHERME DE FREITAS BORGES PEREIRA

**ÉTICA NA UTILIZAÇÃO DE LLM NA PRODUÇÃO
COMPUTACIONAL**

Goiânia
2025

GUILHERME DE FREITAS BORGES PEREIRA

**ÉTICA NA UTILIZAÇÃO DE LLM NA PRODUÇÃO
COMPUTACIONAL**

Monografia e Projeto apresentados ao Curso de Engenharia da Computação da Escola Politécnica da Pontifícia
Universidade Católica de Goiás, para a obtenção do grau de Bacharel em Engenharia da Computação.

Orientador: Me. Claudio Martins Garcia

Goiânia
2025

GUILHERME DE FREITAS BORGES PEREIRA

ÉTICA NA UTILIZAÇÃO DE LLM NA PRODUÇÃO COMPUTACIONAL

Monografia e Projeto apresentados ao Curso de Engenharia da Computação da Escola Politécnica da Pontifícia Universidade Católica de Goiás, para a obtenção do grau de Bacharel em Engenharia da Computação, aprovada em ____ / ____ / ____, pela Banca Examinadora constituída pelos seguintes professores:

Claudio Martins Garcia - orientador
Pontifícia Universidade Católica de Goiás

Dr. Talles Marcelo Gonçalves de Andrade Barbosa
Pontifícia Universidade Católica de Goiás

Me. Pedro Araujo Valle
Pontifícia Universidade Católica de Goiás

Dedico este trabalho a minha mãe,
que foi verdadeiramente paciente com meus defeitos.

AGRADECIMENTOS

A Deus, pela força que me foi dada para ultrapassar os obstáculos no meu caminho ao longo do curso

A minha mãe, que foi verdadeiramente paciente comigo ao longo das muitas falhas na minha jornada.

A minha noiva, que me incentivou a cada passo do caminho.

A meus professores, pelas correções e ensinamentos que me trouxeram ao final deste capítulo na jornada acadêmica.

"Ever tried. Ever failed. No matter. Try again. Fail again. Fail better."

“Tente sempre. Falhe sempre. Não importa. Tente outra vez. Falhe outra vez. Falhe melhor”

Samuel Beckett

RESUMO

Este trabalho busca discutir as implicações éticas decorrentes da utilização de modelos de linguagem de grande escala (LLMs) na produção computacional, com foco em questões como autoria, transparência, viés algorítmico, privacidade e responsabilidade técnica. A partir de uma abordagem qualitativa, foi realizada uma revisão da literatura científica, análise de regulamentações e propostas de instrumentos de coleta de dados, entre eles, um questionário e um roteiro de entrevistas semiestruturadas visando compreender a percepção de profissionais de diferentes áreas sobre o uso ético da inteligência artificial generativa. Embora os dados empíricos ainda não tenham sido coletados, este estudo apresenta uma estrutura metodológica sólida que pode ser aplicada em pesquisas futuras. O objetivo é promover um diálogo informado e responsável sobre a integração dessas tecnologias no cotidiano profissional e social, destacando a necessidade de políticas públicas, educação e regulamentação adequadas. O trabalho visa atingir tanto especialistas quanto o público geral, oferecendo uma reflexão acessível e fundamentada sobre os desafios éticos no uso de LLMs.

Palavras-chave: IA; LLM; Ética; Computação; Produção digital; Inteligência Artificial Generativa.

ABSTRACT

This work addresses the ethical implications arising from the use of large language models (LLMs) in computational content production, focusing on issues such as authorship, transparency, algorithmic bias, privacy, and technical responsibility. Using a qualitative approach, this study includes a review of scientific literature, analysis of regulations, and the proposal of data collection instruments including a questionnaire and semi-structured interviews aimed at understanding professionals' perceptions regarding the ethical use of generative artificial intelligence. Although empirical data has not yet been collected, this research presents a solid methodological framework that can be applied in future studies. The goal is to promote an informed and responsible discussion about the integration of these technologies into professional and social contexts, emphasizing the need for public policies, education, and appropriate regulation. This paper targets both specialists and general audiences, offering a well-founded and accessible reflection on the ethical challenges of using LLMs.

Keywords: AI; LLM; Ethics; Computing; Digital production; Generative Artificial Intelligence.

LISTA DE SIGLAS

LLMs	<i>Large Language Models</i>	12
IA	Inteligencia Artificial	13
AI	<i>Artificial Intelligence</i>	18
LGPD	Lei Geral de Proteção de Dados	23
GDPR	Regulamento Geral sobre a Proteção de Dados	23

LISTA DE SIGLAS.....	9
INTRODUÇÃO.....	12
1. Ética e seu contexto para a computação.....	14
1.1. Conceito clássico de Ética.....	14
1.2. Ética no contexto à computação.....	15
2. Inteligência Artificial, história e impacto.....	17
2.1. O que é Inteligência Artificial.....	17
2.2. Funcionamento de uma inteligência artificial.....	18
2.3. Exemplos de IA no dia a dia.....	18
2.4. Tipos de Inteligência Artificial.....	19
2.5. Um Breve Histórico da Inteligência Artificial.....	19
2.6. Os Pioneiros e os Inícios.....	20
2.7. Os Altos e Baixos.....	20
2.8. A IA Moderna.....	21
3. Revisão da Literatura.....	23
3.1. Contextualização sobre LLMs.....	23
3.2. Questões éticas em modelos de linguagem.....	23
3.3. Viés Algorítmico e Discriminação.....	24
3.4. Privacidade e Consentimento.....	24
3.5. Autoria e Plágio Computacional.....	25
3.6. Transparência e Explicabilidade.....	25
3.7. Responsabilidade Técnica e Jurídica.....	26
3.8. Políticas e Diretrizes Ética Globais.....	26
4. Métodos.....	27
4.1. Pesquisa Bibliográfica.....	27
4.2. Estudo de Casos.....	27
4.3. Análise de Regulamentações e Normas.....	27
4.4. Entrevistas e Questionários.....	28
4.5. Análise e Discussão dos Resultados.....	28
4.6. Proposição de Trabalhos futuros.....	28
5. Resultados e Discussão.....	30
5.1. Conclusões dos materiais estudados.....	30
5.2. Artefatos de captação de dados.....	31
5.2.1. Construção do Questionário e Justificativa das Perguntas.....	31
5.2.2. Entrevista.....	34
5.3. Discussão dos resultados.....	36
6. Conclusão.....	38

REFERÊNCIAS.....	41
APÊNDICE A - Questionario.....	45
APÊNDICE B - Entrevista.....	49

INTRODUÇÃO

Com o crescimento acelerado na utilização e no interesse em modelos de linguagem de grande escala (LLMs, *Large Language Models*) vem sendo cada vez mais relevante a discussão ética da larga adoção e extensa utilização dos mesmos dentro dos processos de automação digital, qualificando-os como ferramenta desses processos e não como produto final.

Desde a disponibilização dos modelos para qualquer usuário e não somente os interessados ou especialistas, vem sendo notado o crescente uso desses modelos para produção digital, em alguns casos, pode se fazer o argumento que muito desse conteúdo é inteiramente dependentes do uso desses LLMs, com nenhuma consideração com a conduta ética no qual este foi produzido.

Este trabalho visa iniciar o diálogo sobre utilização de LLMs como ferramentas e a ética desse uso. Com o objetivo de atingir também o público geral e não só o especialista em tecnologias computacionais e promover um diálogo mais esclarecido sobre o tema, livre de hipérboles.

O assunto foi escolhido como objeto de estudo devido a crescente relevância de modelos de linguagem de grande escala na vida cotidiana do profissional de diversas áreas e também do dia a dia de qualquer usuário padrão de tecnologias da informação, a presença desses modelos é quase inescapável desde a popularização de *chatbots* como *Chat GPT* em meados de 2021. Além disso, a noção do que se configura como uma Inteligência Artificial vem sendo cada vez mais confundida com ficção, devido a popularidade de narrativas da ficção científica como os trabalhos literários de Isaac Asimov, ou filmes como *Blade Runner*.

Foi notado a falta de discussões mais voltadas para os fatos da situação e os desdobramentos da popularização da tecnologia como a dependência nas mesmas para produção intelectual diversa. Este trabalho tem a intenção de conduzir um diálogo a partir da análise aspectos éticos relacionados à utilização de modelos de linguagem de grande escala (LLMs) na produção computacional, investigando seus impactos, desafios e implicações sociais, com o intuito de propor futuras discussões e explorações sobre o tema.

De maneira específica identificar os principais desafios éticos envolvidos na aplicação de LLMs, como viés, privacidade e segurança de dados, além de investigar os impactos sociais e econômicos causados pela implementação de LLMs na produção computacional, incluindo o potencial de substituição de empregos e a responsabilidade social corporativa. Junto dessa investigação, também analisar casos reais de utilização de LLMs na indústria e como esses casos abordam (ou falham em abordar) questões éticas. Avaliar as regulamentações e padrões éticos existentes aplicáveis à utilização de LLMs.

O impacto do objeto de estudo deste trabalho vem sendo cada vez mais um tópico de discussão dentro do círculo tecnológico, de maneira específica, a busca pela IA generalizada acelerou a corrida para a criação da IA verdadeira, o santo graal da automação e o grande sonho da ficção científica. Neste trabalho será proposto um estudo do estado atual do que a indústria chama de Inteligência artificial, seu histórico e sua atuação contemporânea, a fim de gerar um suporte de estudo para discussões futuras sobre o tema, de forma mais objetiva e menos sensacionalista, focado em fatos e dados.

1. Ética e seu contexto para a computação

A ética, enquanto reflexão sobre a conduta humana e os valores que norteiam nossas ações, tem suas raízes na filosofia clássica, especialmente nas obras de Aristóteles, Platão e Sócrates. Segundo Aristóteles, "o homem é por natureza um animal social" (Aristóteles, Política, I, 2), destacando a importância das relações humanas e da responsabilidade individual no convívio coletivo. Essa preocupação com o bem-viver e com as ações orientadas pelo princípio da razão torna-se ainda mais relevante no contexto tecnológico contemporâneo, em que decisões automatizadas impactam diretamente a sociedade. No campo da computação, e especialmente com a ascensão dos modelos de linguagem de grande escala (LLMs), torna-se essencial revisitar os fundamentos éticos para compreender como os avanços tecnológicos podem ser guiados por princípios morais consistentes, garantindo que a inovação ocorra de forma responsável e consciente.

1.1. Conceito clássico de Ética

Na filosofia clássica, a ética transcende os hábitos e costumes meramente estabelecidos pela sociedade. Ela se preocupa em buscar uma fundamentação teórica que permita identificar e praticar o melhor modo de viver e conviver. Essa abordagem implica uma investigação profunda sobre os valores que orientam as ações humanas e as virtudes necessárias para atingir a excelência moral. Para os filósofos antigos, como Aristóteles, a ética estava intrinsecamente ligada ao propósito de alcançar a eudaimonia, ou "felicidade", entendida como a realização plena do potencial humano em harmonia com a sociedade.

Enquanto os costumes representam práticas e normas específicas que variam entre culturas e épocas, a ética clássica propõe princípios universais que transcendem essas particularidades. Sócrates, por exemplo, desafiava constantemente as convenções de seu tempo, questionando se os valores praticados pela sociedade eram realmente bons ou apenas aceitos sem reflexão. Essa postura crítica

inaugurou uma tradição ética que não se satisfaz com respostas prontas, mas exige a investigação racional e o diálogo como caminhos para determinar o que é justo, virtuoso e correto.

Além disso, a ética clássica enfatizava a relação entre o indivíduo e a coletividade. Para Platão, a virtude individual só poderia ser plenamente realizada em uma sociedade bem ordenada, onde cada pessoa desempenhasse seu papel de acordo com suas capacidades. Essa visão aponta para a interdependência entre o bem individual e o bem comum, sugerindo que a ética não é apenas um esforço isolado, mas também um projeto coletivo que exige colaboração e entendimento mútuo.

Assim, a ética na filosofia clássica não se limita a um código de conduta fixo ou a regras imutáveis. Ela é uma disciplina dinâmica que busca compreender a essência do que é viver bem, tanto em nível pessoal quanto em relação à sociedade. Essa busca contínua pela excelência moral permanece relevante até hoje, inspirando reflexões sobre como construir uma vida e uma convivência mais justa e significativa.

1.2. Ética no contexto à computação

No quesito da computação, ética é sensível a contexto de uso, sendo que muitas vezes, a digitalização de um processo está ligada a como esse mesmo pode ser abstraído do mundo real e transformado em um projeto digital, isto é, ética se torna relevante quando falamos do uso do processo digital como ferramenta, meio para um fim.

Um exemplo prático da ética voltada à computação é a produção de código, o ofício de desenvolvedor, de forma geral, sua moralidade é neutra, mas quando especificamente trata de segurança por exemplo, podemos fazer uma análise ética. Abusar de uma vulnerabilidade de segurança pode ser vista como uma ação antiética à primeira vista, mas é amplamente usada na indústria de segurança para

descobrir essa ponta de falha o mais rápido possível com o intuito de correção o mais rápido o possível.

Concluindo o seguinte, a ética no mundo digital deve ser interpretada por aqueles que usam das ferramentas disponíveis, cabe ao usuário o dever ético do bom uso, e não necessariamente a ferramenta em si configura a boa conduta ou a violação ética.

2. Inteligência Artificial, história e impacto

A ideia de máquinas capazes de pensar e agir como seres humanos têm raízes profundas na história da humanidade, mas foi apenas no século XX que os avanços científicos e tecnológicos permitiram dar forma concreta a esse sonho. A Inteligência Artificial (IA) surgiu como um campo científico com a visão de que "máquinas poderiam simular a inteligência humana", proposta oficialmente durante a Conferência de Dartmouth em 1956, considerada o marco inicial da IA moderna. Desde então, a área tem passado por ciclos de otimismo, frustração e renovação, marcados por descobertas revolucionárias e desafios éticos cada vez mais complexos. Este capítulo busca apresentar uma visão geral sobre o conceito, funcionamento e evolução histórica da IA, contextualizando sua relevância nos dias atuais, especialmente com o advento dos modelos de linguagem de grande escala (LLMs).

2.1. O que é Inteligência Artificial

Inteligência Artificial (IA) consiste em um ramo da ciência que visa a criação de um sistema artificial que se assemelha ao pensamento humano, explicitamente seria um sistema ou artefato tecnológico capaz de perceber dados, tratar e raciocinar com base nos mesmos e resolver problemas, isso tudo se assemelhando o máximo possível da razão humana. A inteligência artificial envolve a criação de máquinas inteligentes e não se limita a métodos biologicamente observáveis (MCCARTHY, [s.d.]).

Em específico a noção de Inteligência artificial, apesar de sua interdisciplinaridade, pertence primariamente as áreas computacionais, e consiste na capacidade dos artefatos gerados pela mesma, como *software* ou *hardware*, de ao apresentada um volume de dados, ela possa aprender e se reajustar de acordo com esses dados buscando a melhor resposta disponível a partir dos dados iniciais.

2.2. Funcionamento de uma inteligência artificial

O funcionamento dos grandes modelos de linguagem, considerados a IA moderna, acontece a partir da entrada de grandes volumes de dados, generalizados ou específicos, e por meio do estudo destes dados, usando técnicas próprias para o parcelamento dos mesmos, resultando em um modelo de linguagem natural relativamente capacitado para lidar com entradas humanas e a partir da criação de árvores de decisão, gerar uma resposta esperada.

De forma específica, é necessário mencionar alguns termos quando se fala em inteligência artificial, modelos de linguagem, e a simulação do pensamento humano, sendo esses o aprendizado de máquina e dentro desse contexto temos o *deep learning*.

O aprendizado de máquina é justamente o artefato tecnológico que permite a existência de muitos dos modelos que vemos no mercado hoje. Fazendo uso de aprendizado autônomo e reconhecimentos de padrões, originando se do processamento de grandes volumes de dados, o que permite casamento de padrões incentivando, por meio desses pares, a geração de uma resposta ou raciocínio.

Ainda dentro do contexto de aprendizado de máquina(*machine learning*) temos o conceito de *deep learning*, que como o nome sugere, tem se um aprendizado mais profundo perante os dados apresentados. Fazendo uso de artefatos como redes neurais de alta complexidade, busca um processo análogo ao modelo sináptico do cérebro humano.

2.3. Exemplos de IA no dia a dia

Como dito anteriormente, é quase inescapável a presença de sistemas de inteligência artificial no cotidiano de qualquer indivíduo que tem alguma relação ao mundo digital, dito isso, a IA está presente em diversas áreas da nossa vida, muitas vezes de forma discreta. Alguns exemplos incluem:

- Assistentes virtuais: Siri, Alexa e Google Assistente.
- Recomendação de produtos: Em plataformas de e-commerce e serviços de *streaming*.
- Carros autônomos: Que utilizam sensores e IA para navegar e tomar decisões.
- *Chatbots*: Que atendem clientes em sites e aplicativos.
- Reconhecimento facial: Utilizado em *smartphones* e sistemas de segurança.

2.4. Tipos de Inteligência Artificial

Existem diferentes tipos de IA, que podem ser classificados de acordo com sua complexidade e capacidade:

- A Fraca (Narrow AI): São sistemas desenvolvidos para realizar tarefas específicas, como jogar xadrez, traduzir idiomas ou recomendar conteúdos. Atualmente, todas as formas de IA em operação pertencem a essa categoria (RUSSELL; NORVIG, 2020).
- IA Forte (General AI): Representa uma inteligência artificial hipotética, capaz de compreender, aprender e aplicar conhecimento em qualquer contexto intelectual — algo até agora inexistente na prática. Sua criação levantaria profundas questões filosóficas e éticas, tema central de debates como os apresentados por Bostrom (2014).

Em resumo, a Inteligência Artificial é uma área em constante evolução que está transformando a maneira como vivemos e trabalhamos

2.5. Um Breve Histórico da Inteligência Artificial

A Inteligência Artificial (IA) não é um conceito novo. A ideia de criar máquinas capazes de pensar e aprender como humanos fascina a humanidade há séculos. No

entanto, foi apenas no século XX que os avanços tecnológicos permitiram dar os primeiros passos concretos nessa direção.

2.6. Os Pioneiros e os Inícios

Os fundamentos teóricos da Inteligência Artificial (IA) começaram a ser estabelecidos ainda no início do século XX, impulsionados pelo avanço da lógica matemática e das ciências cognitivas. Um dos marcos iniciais foi a introdução do conceito de "Máquina de Turing", proposto por Alan Turing em 1936. Esse modelo teórico demonstrou como uma máquina poderia executar qualquer tarefa computável, servindo como base para o desenvolvimento posterior de sistemas inteligentes.

O termo "Inteligência Artificial", porém, só surgiu oficialmente em 1956, durante a Conferência de Dartmouth, organizada por John McCarthy, Marvin Minsky, Nathaniel Rochester e Claude Shannon. Este evento é considerado o marco inicial da IA como área de estudo formal. Durante esse encontro, os pesquisadores expressaram otimismo quanto à possibilidade de máquinas aprenderem e resolverem problemas complexos de forma autônoma.

Nos anos seguintes, surgiram os primeiros programas capazes de simular raciocínio humano, como o *Logic Theorist*, criado por Allen Newell e Herbert Simon, que provava teoremas matemáticos, e o *General Problem Solver*, que buscava solucionar problemas de maneira mais ampla. Esses desenvolvimentos deram início ao período conhecido como "verão da IA".

2.7. Os Altos e Baixos

Apesar do entusiasmo inicial, as expectativas geradas nos primeiros anos da IA não foram totalmente cumpridas. Nos anos 1970, a falta de avanços significativos na capacidade real das máquinas e os altos custos computacionais levaram ao chamado "Inverno da IA", um período caracterizado pela redução de investimentos e interesse institucional na área.

No entanto, a partir dos anos 1980, a IA voltou a ganhar força com o desenvolvimento de sistemas especialistas , softwares capazes de simular o conhecimento e o raciocínio de especialistas humanos em áreas específicas. Essas ferramentas eram amplamente utilizadas em diagnósticos médicos, análise financeira e outras aplicações técnicas. Apesar do sucesso inicial, esses sistemas revelaram-se limitados em sua capacidade de adaptação e generalização, levando novamente a uma desaceleração no progresso.

Esse ciclo de altos e baixos reflete a complexidade do problema da criação de sistemas realmente inteligentes, exigindo avanços simultâneos em *hardware*, algoritmos e compreensão do funcionamento da mente humana.

2.8. A IA Moderna

A partir dos anos 1990, a IA passou por uma nova fase de desenvolvimento, impulsionada principalmente pelo crescimento exponencial da capacidade computacional e pela disponibilidade de grandes volumes de dados — o que deu origem ao conceito de *Big Data* .

Um dos grandes avanços dessa fase foi o fortalecimento do aprendizado de máquina (*machine learning*), uma subárea da IA que permite aos sistemas aprenderem a partir de dados sem serem explicitamente programados. Métodos como redes neurais artificiais, árvores de decisão e máquinas de vetores de suporte tornaram-se populares nesse contexto.

Nos últimos anos, o *deep learning* (aprendizado profundo) tem revolucionado o campo da IA. Baseado em redes neurais artificiais com múltiplas camadas, o *deep learning* permitiu avanços significativos em áreas como visão computacional, processamento de linguagem natural e reconhecimento de padrões. Modelos como AlexNet , ResNet , BERT e GPT marcaram momentos importantes na evolução da IA moderna.

Além disso, a popularização de frameworks como *TensorFlow* e *PyTorch* e a disponibilização de conjuntos de dados públicos fizeram com que a IA deixasse de ser exclusividade de laboratórios especializados e se tornasse uma tecnologia amplamente acessível, integrando produtos e serviços do cotidiano.

3. Revisão da Literatura

De forma a compreender mais profundamente as implicações éticas associadas ao uso de modelos de linguagem de grande escala (LLMs) na produção computacional, é fundamental revisitar estudos e pesquisas já realizados sobre inteligência artificial, seus desafios éticos e suas aplicações práticas. O presente capítulo tem como objetivo apresentar uma síntese crítica da literatura relevante, abordando conceitos teóricos, debates atuais e diretrizes normativas que fundamentam a discussão ética em torno da utilização de LLMs. Serão explorados temas como viés algorítmico, privacidade, autoria, transparência e responsabilidade técnica, com base em artigos científicos, relatórios técnicos e políticas públicas nacionais e internacionais.

3.1. Contextualização sobre LLMs

Os *Large Language Models* (LLMs) são sistemas computacionais treinados com grandes volumes de dados textuais, capazes de compreender e gerar linguagem natural de forma autônoma. Modelos como GPT-3 (Brown et al., 2020), BERT (Devlin et al., 2018) e LLaMA (Meta, 2023) representam avanços significativos na capacidade de máquinas em simular processos cognitivos humanos, especialmente no que diz respeito à escrita e interpretação textual.

Esses modelos têm sido amplamente utilizados em diversos contextos: desde assistentes virtuais até geração automatizada de conteúdo acadêmico, jornalístico e criativo. No entanto, essa ampla aplicação levanta questionamentos importantes sobre as implicações éticas do uso dessas tecnologias, especialmente quando empregadas de forma pouco crítica ou sem supervisão humana.

3.2. Questões éticas em modelos de linguagem

De acordo com Bender et al. (2021), o aumento exponencial do tamanho dos modelos de linguagem traz consigo riscos éticos significativos. Esses autores alertam para o fato de que modelos maiores não necessariamente são melhores,

especialmente quando consideramos impactos sociais, ambientais e culturais. Entre esses riscos, destacam-se:

- Viés algorítmico : LLMs absorvem padrões discriminatórios presentes nos dados de treinamento.
- Falta de transparência : Muitas vezes, os mecanismos internos de decisão são opacos tanto para usuários quanto para desenvolvedores.
- Impacto ambiental : O custo energético para treinar modelos é elevado.
- Questões de autoria e propriedade intelectual : Conteúdo gerado por AI levanta dúvidas sobre quem deve ser creditado.

3.3. Viés Algorítmico e Discriminação

Estudos mostram que LLMs podem reproduzir preconceitos existentes na sociedade, como viés racial, de gênero e de classe (Buolamwini & Gebru, 2018; Zhao et al., 2017). Isso ocorre porque os modelos aprendem a partir de grandes conjuntos de dados coletados da internet, onde há desequilíbrios estatísticos e tendências culturais dominantes.

Por exemplo, pesquisas apontam que modelos de linguagem tendem a associar palavras relacionadas a profissões com gêneros específicos (Baker et al., 2018). Essas associações podem reforçar estereótipos e perpetuar desigualdades se não forem devidamente controladas.

3.4. Privacidade e Consentimento

A utilização de dados pessoais para treinamento de LLMs também levanta preocupações sobre privacidade e consentimento informado. A maioria dos modelos modernos é treinada com textos coletados de domínios públicos, redes sociais, livros e artigos científicos — muitas vezes sem que os autores originais saibam ou concordem com esse uso.

O Regulamento Geral de Proteção de Dados (GDPR) da União Europeia e a Lei Geral de Proteção de Dados (LGPD) no Brasil buscam proteger os direitos individuais nesse contexto, mas ainda há lacunas na regulamentação específica para modelos de linguagem.

3.5. Autoria e Plágio Computacional

Uma das questões mais discutidas recentemente é a referente à autoria do conteúdo produzido por LLMs. Quando um modelo gera um texto, ele está replicando padrões aprendidos durante o treinamento, mas pode incluir informações muito próximas ou idênticas a conteúdos previamente existentes. Isso levanta debates sobre plágio, originalidade e direitos autorais (Ienca et al., 2018).

Além disso, há casos em que estudantes e profissionais utilizam LLMs para escrever trabalhos acadêmicos ou conteúdos comerciais sem atribuição adequada, o que pode configurar má-fé acadêmica ou violação de normas editoriais.

3.6. Transparência e Explicabilidade

Um dos grandes desafios na adoção ética de LLMs é a falta de transparência em relação ao funcionamento interno dos modelos. Mittelstadt et al. (2016) destacam que a complexidade dos sistemas de IA dificulta a explicação clara de como certas decisões ou outputs são gerados.

Para mitigar esse problema, pesquisadores sugerem práticas como:

- Documentação detalhada do processo de treinamento;
- Publicação de relatórios técnicos e éticos pelos desenvolvedores;

- Uso de métricas de justiça algorítmica para avaliar os modelos antes da implantação

3.7. Responsabilidade Técnica e Jurídica

Com o aumento da dependência de sistemas automatizados, surge a questão de quem é responsável por decisões tomadas ou conteúdos gerados por LLMs. Binns (2018) argumenta que a responsabilidade deve ser compartilhada entre desenvolvedores, empresas e usuários finais.

No âmbito jurídico, há poucas leis específicas que abordem diretamente a responsabilização em casos de danos causados por IA. Enquanto isso, recomendações éticas sugerem a criação de mecanismos de auditoria e controle prévio à liberação de novos modelos ao público.

3.8. Políticas e Diretrizes Ética Globais

Diversas organizações internacionais têm lançado diretrizes para orientar o uso ético da inteligência artificial. A UNESCO (2021), por exemplo, publicou uma recomendação global sobre ética em IA, enfatizando princípios como:

- Respeito aos direitos humanos
- Justiça social
- Inclusão e diversidade
- Sustentabilidade ambiental
- Governança participativa

Da mesma forma, a OCDE (2019) e a União Europeia (*Ethics Guidelines for Trustworthy AI*, 2019) apresentaram marcos normativos para garantir que a IA seja segura, transparente e responsável.

4. Métodos

A metodologia adotada neste trabalho envolve uma abordagem qualitativa e descritiva, com o objetivo de explorar e analisar as questões éticas relacionadas ao uso de modelos de linguagem de grande escala (LLMs) na produção computacional. A seguir, são descritas as etapas que serão seguidas

4.1. Pesquisa Bibliográfica

Será realizada uma revisão de literatura abrangente, focada em estudos acadêmicos, artigos científicos, livros e relatórios técnicos sobre LLMs, inteligência artificial, e ética na computação. A revisão terá o intuito de contextualizar os principais conceitos, identificar os desafios éticos e examinar as abordagens propostas para enfrentar tais questões.

As fontes serão selecionadas por meio de bases de dados como *IEEE Xplore*, *Google Scholar*, *ACM Digital Library*, e *Scopus*, buscando materiais publicados nos últimos 10 anos para garantir relevância e atualidade.

4.2. Estudo de Casos

Serão analisados estudos de casos reais de empresas e organizações que utilizam LLMs em processos de produção computacional, com foco em exemplos que demonstrem boas práticas ou problemas éticos enfrentados.

O estudo de casos possibilitará a análise crítica da aplicação prática de LLMs, investigando como esses casos abordam questões como viés, privacidade de dados, transparência, e responsabilidade.

4.3. Análise de Regulamentações e Normas

A metodologia incluirá uma análise comparativa das principais regulamentações e normas éticas aplicáveis ao uso de inteligência artificial e LLMs. Serão analisados documentos como o Regulamento Geral de Proteção de Dados (GDPR), iniciativas da União Europeia e da Organização para a Cooperação e Desenvolvimento Econômico (OCDE) e outros marcos legais relevantes.

O objetivo é identificar lacunas normativas e pontos onde a legislação pode ser aprimorada para acompanhar o desenvolvimento tecnológico.

4.4. Entrevistas e Questionários

Serão realizados questionários e entrevistas com profissionais de diversas áreas, a fim de obter *insights* e opiniões sobre as práticas éticas no uso de LLMs.

As entrevistas serão estruturadas e realizadas de forma online, seguindo um roteiro pré-definido que aborda temas como privacidade, viés algorítmico e transparência.

4.5. Análise e Discussão dos Resultados

Os dados coletados na pesquisa bibliográfica, estudo de casos, e possíveis entrevistas serão analisados qualitativamente para identificar padrões, convergências e divergências em relação aos desafios éticos no uso de LLMs.

A análise será feita à luz de teorias éticas, como a deontologia, utilitarismo e ética de virtude, aplicadas ao contexto tecnológico.

4.6. Proposição de Trabalhos futuros

A partir da análise crítica do material estudado, serão elaboradas recomendações e estruturas de apoio para trabalhos futuros que, usando artefatos e moldes presentes nesses trabalhos, poderão aprofundar no estudo ético e prático do impacto do uso de Inteligência artificial na produção digital contemporânea.

5. Resultados e Discussão

Com base na análise bibliográfica, estudo histórico e contextualização contemporânea, este capítulo abrange a formação de artefatos de suporte à pesquisa e captação de dados para o estudo de forma prática e aplicada dos impactos éticos na produção digital quando envolvida com o uso de Inteligência Artificial.

5.1. Conclusões dos materiais estudados

A partir da revisão da literatura e da análise das bases teóricas sobre ética e inteligência artificial, é possível identificar que a utilização de modelos de linguagem de grande escala (LLMs) na produção computacional traz consigo um conjunto significativo de implicações éticas que não podem ser ignoradas.

Esses modelos, embora amplamente difundidos por sua capacidade de gerar textos coerentes e em tempo real, operam com base em dados coletados de fontes diversas, muitas vezes sem consentimento explícito dos autores originais, levantando preocupações quanto à privacidade, autoria e direitos autorais.

Um dos principais aspectos discutidos foi o viés algorítmico, isto é, a tendência dos modelos de reproduzir preconceitos existentes na sociedade, como desigualdades de gênero, etnia e classe social.

Esse fenômeno pode impactar negativamente decisões automatizadas em contextos críticos, como recrutamento profissional, sistemas educacionais e até mesmo assistência médica, quando os LLMs são usados como suporte à tomada de decisão.

Outro ponto central abordado foi a questão da transparência e explicabilidade. Muitos LLMs atuam como "caixas pretas", ou seja, seus processos internos de decisão são opacos tanto para usuários finais quanto para pesquisadores.

Isso compromete a possibilidade de auditoria e responsabilização em casos de falhas ou danos causados pelo uso inadequado desses sistemas.

Quanto às políticas e regulamentações , verificou-se que há avanços importantes em nível global, como o Regulamento Geral de Proteção de Dados (GDPR), as diretrizes da OCDE e a Recomendação da UNESCO sobre ética da IA.

No entanto, ainda persistem lacunas normativas específicas para modelos de linguagem, especialmente no Brasil, onde a Lei Geral de Proteção de Dados (LGPD) oferece algum respaldo, mas carece de atualizações para lidar com os novos desafios impostos pela IA generativa.

Por fim, constata-se que a ética no uso de LLMs não deve ser encarada apenas como uma preocupação técnica ou jurídica, mas como uma responsabilidade coletiva que envolve desenvolvedores, empresas, usuários e formuladores de políticas.

A educação e conscientização sobre os riscos e limitações dessas tecnologias são fundamentais para promover um uso mais responsável e consciente.

5.2. Artefatos de captação de dados

Ao observar a crescente tendência de adoção em massa da IA no processo de produção digital fica claro a necessidade do diálogo mais formalizado ancorado em dados e não em noções que beiram a ficção, em juízo disso um estudo tanto quantitativo quanto qualitativo é proposto para analisar o impacto dessa adesão em diversas áreas.

5.2.1. Construção do Questionário e Justificativa das Perguntas

A partir do objetivo qualitativo do estudo, foi elaborado um questionário com o intuito de compreender a percepção de pessoas de diferentes áreas sobre os desafios éticos associados ao uso de inteligência artificial na produção digital.

Embora ainda não tenha sido aplicado, o questionário foi estruturado para capturar informações sobre familiaridade com IA, visões sobre regulação, impacto no mercado de trabalho, transparência, autoria, viés algorítmico, responsabilidade técnica e aceitação do avanço tecnológico.

Cada uma das perguntas foi pensada para permitir análises posteriores mais profundas, tanto quantitativas quanto qualitativas, caso sejam utilizadas em pesquisas futuras. O questionário modelo referente ao anexo A.

A primeira pergunta busca identificar o perfil profissional do respondente, apresentando opções como Educação, Tecnologia da Informação (TI), Direito, Medicina ou outra área a ser especificada. Esse dado é fundamental para contextualizar as respostas, considerando que diferentes campos profissionais possuem distintas interações e perspectivas sobre a utilização de IA generativa.

A segunda pergunta investiga o grau de contato do respondente com ferramentas de IA, oferecendo alternativas que variam desde uso frequente até desconhecimento total sobre essas tecnologias. A expectativa é que essa informação ajude a entender como a experiência prática influencia a percepção sobre ética e responsabilidade no uso dessas ferramentas.

A terceira pergunta aborda a necessidade de regulamentação específica para o uso de IA na produção digital, com respostas que vão desde apoio a regras claras e específicas até a defesa da autorregulação do setor. Essa questão visa explorar a demanda social por normatização e a confiança nos mecanismos atuais de controle.

A quarta pergunta procura avaliar a visão do respondente sobre a relação entre IA e direitos autorais, questionando se o uso de modelos de linguagem afeta significativamente a autoria de conteúdos criados por humanos. As alternativas permitem refletir sobre a IA como ferramenta de apoio ou como potencial substituta na criação original.

A quinta pergunta explora o impacto da IA no mercado de trabalho dentro da área de atuação do respondente, com opções que vão desde substituição completa de profissionais até reconhecimento de pouca influência. O objetivo é mapear a percepção sobre automação e seus efeitos sociais e econômicos.

A sexta pergunta levanta a questão da transparência, indagando se é adequado criar conteúdos com IA sem informar que foram gerados por uma máquina. As alternativas buscam compreender até que ponto a honestidade informativa é vista como obrigatória ou opcional em diferentes contextos.

A sétima pergunta analisa a percepção de risco relacionada ao uso de IA na geração de conteúdos enganosos, como desinformação e *deepfakes*. Com alternativas que vão desde reconhecimento de gravidade até negação do problema, essa questão busca medir o nível de preocupação com a manipulação e falsificação digital.

A oitava pergunta investiga a crença de que a IA pode reproduzir ou reforçar preconceitos existentes na sociedade, com destaque para a origem dos dados de treinamento e a possibilidade de mitigação através de boas práticas técnicas. É uma forma de abordar o tema do viés algorítmico e sua percepção pública.

A nona pergunta busca entender quem deve ser considerado responsável legalmente por danos causados por conteúdos gerados por IA, com opções que envolvem usuário, desenvolvedor ou compartilhamento da responsabilidade. Este item é relevante para discussões futuras sobre marcos legais e fiscalização.

A décima pergunta examina a relação entre IA e criatividade humana, questionando se a utilização de modelos de linguagem limita ou estimula a produção criativa. Essa reflexão é importante para áreas como arte, escrita e design, onde a originalidade é central.

A décima primeira pergunta avalia o grau de confiança nas decisões tomadas por sistemas de IA em diferentes áreas profissionais, destacando a importância da supervisão humana mesmo em contextos altamente automatizados.

A décima segunda pergunta traz uma visão prospectiva, perguntando se o avanço da IA na produção digital deve continuar sendo incentivado, com alternativas que consideram liberdade total, acompanhamento regulatório ou contenção estratégica.

Por fim, a décima terceira pergunta propõe um cenário hipotético, solicitando ao respondente que imagine acesso a uma IA altamente avançada, porém com margem de erro conhecida, perguntando qual seria o limite aceitável para confiar em suas respostas. Essa questão visa

mensurar a tolerância ao risco e a percepção de confiabilidade em sistemas de IA.

5.2.2. Entrevista

As entrevistas são uma ferramenta metodológica valiosa para aprofundar temas complexos e obter percepções qualitativas sobre questões éticas no uso de modelos de linguagem de grande escala (LLMs).

Diferentemente do questionário fechado, que permite maior abrangência quantitativa, as entrevistas possibilitam explorar nuances, experiências pessoais e reflexões críticas que enriquecem a análise do estudo.

Para garantir a qualidade e relevância das entrevistas, é essencial selecionar participantes com conhecimento direto ou indireto sobre os temas centrais do trabalho: ética, inteligência artificial e uso de LLMs na produção digital. Sendo os critérios para seleção de entrevistados:

- Especialistas em IA: Profissionais da área de tecnologia com experiência prática em desenvolvimento ou implementação de modelos de linguagem.
- Profissionais de áreas regulatórias ou jurídicas: Advogados, especialistas em direito digital ou proteção de dados, familiarizados com as implicações legais do uso de IA.
- Acadêmicos e pesquisadores: Professores ou estudiosos com publicações em áreas como ética aplicada à tecnologia, filosofia da ciência ou IA.
- Profissionais criativos: Jornalistas, designers, escritores ou artistas que já utilizaram ou têm opinião formada sobre o impacto da IA na criação humana.

- Representantes de órgãos governamentais ou instituições públicas: Aqueles envolvidos na formulação de políticas públicas relacionadas à IA ou educação digital.

Após a etapa de seleção dos entrevistados, vem a consideração dos questionamentos a serem discutidos para a captação de dados e nuances relevantes à discussão ética e tecnológica do tema. As entrevistas seguiram um roteiro semiestruturado, ou seja, com perguntas previamente definidas, mas permitindo liberdade ao entrevistado para expandir suas respostas.

O objetivo é obter tanto informações objetivas quanto reflexões profundas sobre o tema. Exemplo de roteiro de entrevista encontrado no anexo B

Como observações adicionais sobre a etapa de entrevistas, as mesmas poderão ser realizadas presencialmente ou virtualmente, mediante prévia autorização do entrevistado para gravação de áudio ou vídeo, caso seja necessário para a fidelidade das informações coletadas.

Para a análise das respostas, será utilizado o método de análise temática, com o objetivo de identificar categorias e extrair insights relevantes que contribuam para a compreensão do tema estudado.

Todos os dados obtidos serão tratados com confidencialidade e, no relatório final, serão apresentados de forma anonimizada, garantindo o sigilo e a privacidade dos participantes.

5.3. Discussão dos resultados

A partir do objetivo qualitativo deste trabalho, foi elaborada uma proposta de coleta de dados por meio de questionário e entrevistas semiestruturadas com o intuito de aprofundar a compreensão sobre a percepção da sociedade frente ao uso ético de modelos de linguagem de grande escala (LLMs) na produção digital.

Embora esses instrumentos ainda não tenham sido aplicados, sua estrutura foi desenvolvida visando capturar opiniões de indivíduos de diferentes áreas profissionais, permitindo assim uma análise multidisciplinar e representativa.

O questionário foi elaborado com questões fechadas e uma questão aberta final, buscando equilibrar a objetividade dos dados com a profundidade interpretativa. Ele aborda temas como familiaridade com ferramentas de IA, percepção sobre autoria e direitos autorais, impacto no mercado de trabalho, transparência no uso da tecnologia, viés algorítmico, responsabilidade jurídica e aceitação do avanço tecnológico.

A expectativa é que, ao ser aplicado, ele forneça um retrato quantitativo da aceitação e preocupação da sociedade diante do uso crescente dessas tecnologias, especialmente em áreas criativas e técnicas.

Já as entrevistas, previstas como semiestruturadas, têm como foco principal aprofundar os aspectos qualitativos da pesquisa, permitindo a exploração de nuances que o questionário sozinho não consegue capturar. Serão direcionadas a especialistas de diferentes áreas como tecnologia, direito, educação e comunicação com o objetivo de obter reflexões críticas e fundamentadas sobre os desafios éticos envolvendo a IA generativa.

O roteiro previsto busca explorar conceitos como autonomia humana frente à automação, necessidade de regulamentação específica, impactos na criatividade e autoria, além da importância da transparência e da educação nesse contexto.

6. Conclusão

A escolha por abordar a ética no uso de modelos de linguagem de grande escala (LLMs) partiu da necessidade de ultrapassar o viés meramente tecnológico e examinar a realidade de aplicação dessas ferramentas no cotidiano das pessoas, seja no ambiente profissional ou pessoal. Diferentemente de um estudo puramente técnico ou de desenvolvimento de algoritmos, este trabalho buscou compreender os impactos sociais, culturais e éticos trazidos pela popularização de sistemas de IA generativa, especialmente em contextos onde a supervisão humana é mínima ou inexistente.

Ao longo da pesquisa, percebeu-se que há uma crescente dependência de ferramentas baseadas em IA, como chatbots e assistentes de escrita, que são amplamente utilizados na produção de textos, programação, criação artística e até mesmo na elaboração de conteúdos acadêmicos. No entanto, muitos usuários empregam essas ferramentas sem questionar a origem dos dados, os vieses incorporados ou as implicações legais e morais envolvidas. Isso evidencia uma lacuna entre a facilidade de uso e o entendimento crítico sobre os limites, responsabilidades e consequências do uso de LLMs.

Uma das principais reflexões desenvolvidas ao longo deste estudo é a de que os chamados “modelos de inteligência artificial generativa” não possuem, de fato, uma capacidade criativa no sentido humano, mas sim uma habilidade sofisticada de reconhecer padrões e gerar respostas com base em grandes volumes de informações previamente absorvidas. Em termos mais precisos, eles operam como sistemas avançados de associação e previsão de palavras, sendo mais próximos de um “*auto-complete* inteligente” do que de uma inteligência propriamente dita. Essa distinção é fundamental para evitar a personificação dessas máquinas e manter o debate focado no papel humano na condução e supervisão da tecnologia.

Com isso, surge a necessidade de avaliar criticamente o uso desses modelos sob uma perspectiva ética clássica, considerando princípios como justiça, responsabilidade, transparência e o bem-comum. Questões como autoria, direitos autorais, consentimento no uso de dados e viés algorítmico foram destacadas ao longo do trabalho como desafios centrais, muitas vezes negligenciados tanto pelos usuários quanto pelos desenvolvedores. Afinal, se o conteúdo produzido por um LLM pode ser resultado de treinamento feito com obras de autores não consultados ou remunerados, podemos afirmar que esse modelo realmente "cria", ou apenas recombina?

O escopo deste trabalho foi deliberadamente voltado para a análise conceitual, histórica e ética, com o objetivo de promover uma discussão acessível e fundamentada, capaz de atingir tanto especialistas quanto o público geral. Embora tenham sido propostos instrumentos de coleta de dados, como questionário e roteiro de entrevista, o estudo permanece, por ora, no campo qualitativo e teórico, servindo como base para futuras investigações mais aprofundadas e empíricas.

Como contribuição imediata, este trabalho busca incentivar um diálogo mais responsável sobre o uso de LLMs na produção digital, rompendo com narrativas sensacionalistas ou exageradas sobre a capacidade da IA e estimulando uma reflexão madura e consciente sobre suas reais implicações. Além disso, pretende-se oferecer suporte para debates interdisciplinares, integrando áreas como filosofia, direito, ciência da computação e comunicação, já que os dilemas éticos decorrentes do uso de IA transcendem o domínio técnico e exigem posicionamentos sociais e institucionais sólidos.

Para trabalhos futuros, sugerem-se:

- A aplicação prática do questionário elaborado (Apêndice A), visando mapear percepções de diferentes perfis profissionais sobre o uso ético de LLMs;
- A realização de entrevistas semiestruturadas com especialistas em IA, juristas, educadores e criadores digitais, a fim de capturar nuances qualitativas sobre dilemas éticos no uso da IA;

- A análise de casos concretos de uso de LLMs em ambientes reais, como empresas de mídia, instituições de ensino e plataformas de conteúdo, para identificar boas práticas e falhas recorrentes;
- A exploração de propostas regulatórias e normativas, tanto no Brasil quanto internacionalmente, para guiar o uso responsável dessa tecnologia;
- E, por fim, a promoção de debates multidisciplinares que integrem tecnologia, ética e educação, preparando melhor a sociedade para lidar com os desafios e oportunidades trazidos pela revolução da IA generativa.

Em síntese, este trabalho apresenta-se como um primeiro passo para pensar a inteligência artificial não apenas como uma ferramenta técnica, mas como um fenômeno social que demanda atenção constante, formação crítica e regulação adequada. À medida que os LLMs se tornam cada vez mais presentes em nossas vidas, urge refletir: não apenas como usamos a IA, mas por quê, para quem e com quais consequências.

REFERÊNCIAS

ARISTÓTELES . *Política* . Tradução de Roberto Leal Ferreira. São Paulo: Nova Cultural, 1997. (Coleção Os Pensadores).

[«O que são grandes modelos de linguagem? — Explicação sobre a IA do LLM — AWS»](#). Amazon Web Services, Inc. Consultado em 25 de novembro de 2024

Neelakandan, Laya (26 de abril de 2023). [«O que são grandes modelos de linguagem \(LLMs\) e para que servem»](#). Fast Company Brasil. Consultado em 25 de novembro de 2024

MCCARTHY, John. *Artificial Intelligence: What is AI?* Disponível em: <http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html>. Acessado em 26 de novembro 2024

MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N.; SHANNON, C. E. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine* , v. 27, n. 4, p. 12–14, 2006. Disponível em: <https://ojs.aaai.org/aimagazine/index.php/aimagazine/article/view/1904>. Acesso em: 5 out. 2024.

SOARES Luana Murito Bellinello; PIRES, Rosiléa dos Santos Amatto. Reflexões acerca do imediatismo generalizado no espaço escolar. *Revista Tecnologia Educacional* [on line], Rio de Janeiro, n. 236, p.69-83, 2023. ISSN: 0102-5503. Consultado em 26 de novembro 2024

BAKER, T. et al. Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings . In: *Advances in Neural Information Processing Systems*, v. 29, 2016.

BENDER, E. M. et al. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, New York, NY, USA, 2021. p. 610–623.

BINNS, R. *Algorithmic Accountability and Transparency in Machine Learning* . PhD Thesis, University of Cambridge, 2018.

BRUNDATE, P. *et al.* *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* . arXiv:1802.07228 [cs.CY], 2018.

BUOLAMWINI, J.; GEBOURG, T. *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification* . In: *Proceedings of the Conference on Fairness, Accountability and Transparency – FAT**, Nova York, 2018. p. 77–91.

DEVLIN, J. *et al.* *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* . In: *NAACL HLT*, 2019. Disponível em: <https://arxiv.org/abs/1810.04805> . Consultado em 20 de maio 2025

FLORIDI, L. *et al.* *AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations* . *Minds and Machines*, v. 28, p. 689–707, 2018.

GARRETT, M. *Who Owns AI-Generated Content? Copyright Challenges in the Age of Generative AI* . *Harvard Journal of Law & Technology*, v. 36, n. 1, 2023.

GENERAL DATA PROTECTION REGULATION (GDPR). *Regulation (EU) 2016/679. Official Journal of the European Union*, 2016.

IENCA, M. *et al.* *Towards a Human Right to Ethical Design of Brain-Computer Interfaces* . *Science and Engineering Ethics*, v. 24, p. 1703–1725, 2018.

LGPD – Lei Geral de Proteção de Dados (Brasil). Lei nº 13.709, de 14 de agosto de 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm . Acessado em 21 de maio 2025

MITTELSTADT, B. *et al.* *The Ethics of Algorithms: Mapping the Debate* . *Big Data & Society*, v. 3, n. 2, p. 1–21, 2016.

NISSENBAUM, H. *Privacy as Contextual Integrity* . *Washington Law Review*, v. 79, n. 1, p. 119–158, 2004.

OECD. *OECD AI Principles* . *Organisation for Economic Co-operation and Development*, 2019. Disponível em: <https://www.oecd.org/en/topics/sub-issues/ai-principles.html>. Acessado em 18 de maio de 2025

RAFFEL, C. *et al.* *Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer* . *Journal of Machine Learning Research*, v. 21, 2020.

UNESCO. *Recommendation on the Ethics of Artificial Intelligence* . Paris, 2021. Disponível em: <https://www.unesco.org/en/articles/recommendation-ethics-artificial-intelligence>

VASWANI, A. *et al.* *Attention Is All You Need* . In: *Advances in Neural Information Processing Systems*, v. 30, 2017.

ZERILLI, J. *et al.* *Transparency in Algorithmic Decision-Making: Opportunities and Challenges*. *Communications of the ACM*, v. 61, n. 9, p. 84–91, 2018.

GOOGLE. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*, 2015. Disponível em: <https://www.tensorflow.org/> . Acesso em: 10 maio 2025.

FAIR (Facebook AI Research). *PyTorch* . Disponível em: <https://pytorch.org/> . Acesso em: 10 maio 2025.

TURING, A. M. *On Computable Numbers, with an Application to the Entscheidungsproblem* . *Proceedings of the London Mathematical Society*, v. 42, p. 230–265, 1936. Disponível em: <https://doi.org/10.1112/plms/s2-42.1.230> . Acesso em: 20 maio 2025.

NEWELL, A.; SHAW, J. C.; SIMON, H. A. *Elements of a Theory of Human Problem Solving*. *Psychological Review* , v. 65, n. 3, p. 151–166, 1958.

NEWELL, A.; SIMON, H. A. Human Problem Solving: The State of the Theory in Problem Solving Research. In: Proceedings of the Seventh International Congress on Logic, Methodology and Philosophy of Science , 1973.

NEWELL, A.; SHAW, J. C.; SIMON, H. A. Report on a General Problem-Solving Program. In: Proceedings of the International Conference on Information Processing , p. 256–264, 1959.

RUSSELL, S.; NORVIG, P. Artificial Intelligence: A Modern Approach . 4. ed. Pearson, 2020.

BOSTROM, N. Superintelligence: Paths, Dangers, Strategies . Oxford University Press, 2014.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning . MIT Press, 2016.
Disponível em: <http://www.deeplearningbook.org/>. Acesso em: 06 junho 2025.

APÊNDICE A - Questionário

Contextualização:

A Inteligência Artificial (IA) é um conjunto de tecnologias capazes de simular comportamentos humanos, como reconhecer padrões, aprender com dados e tomar decisões. Atualmente, ela está presente em muitas áreas do cotidiano — desde recomendações em redes sociais até ferramentas que criam textos, imagens, vídeos ou códigos automaticamente.

Na produção digital (design, marketing, jornalismo, audiovisual, etc.), a IA tem sido usada para gerar conteúdos de forma rápida e acessível. Isso traz benefícios, mas também levanta dúvidas sobre autoria, transparência, impactos no mercado de trabalho e uso responsável da tecnologia.

Este questionário tem como objetivo compreender a percepção de pessoas de diferentes áreas sobre essas questões. As respostas são anônimas e serão utilizadas exclusivamente para fins acadêmicos, no contexto de um Trabalho de Conclusão de Curso (TCC).

Instruções:

Por favor, leia cada pergunta com atenção e marque a alternativa que melhor representa a sua opinião. Caso não se sinta confortável ou seguro(a) para responder, há a opção "Não me sinto apto(a) a responder".

1. Qual é a sua área de atuação?

Educação

Tecnologia da Informação (TI)

Direito

Medicina

Outra: _____

2. Você já utilizou ferramentas de IA para criar ou modificar conteúdos digitais (como textos, imagens, vídeos, músicas ou códigos)?

- Sim, com frequência
- Sim, algumas vezes
- Não, mas conheço essas ferramentas
- Não, e não conheço essas ferramentas
- Não me sinto apto(a) a responder

3. Na sua opinião, o uso de IA na produção digital deveria ser regulamentado?

- Sim, com regras claras e específicas
- Sim, mas de forma flexível e adaptável
- Não, apenas orientações éticas são suficientes
- Não, o setor deve se autorregular
- Não me sinto apto(a) a responder

4. Você acredita que o uso de IA pode afetar os direitos autorais e a autoria de conteúdos criados por humanos?

- Sim, de forma significativa
- Sim, em alguns casos
- Não, se houver transparência no uso da IA
- Não, a IA é apenas uma ferramenta de apoio
- Não me sinto apto(a) a responder

5. Em sua área, você acredita que a IA pode influenciar a necessidade de trabalho humano?

- Sim, substituindo alguns profissionais
- Sim, mas apenas em tarefas repetitivas
- Não, a IA complementa o trabalho humano
- Não, a IA tem pouco impacto na área
- Não me sinto apto(a) a responder

6. Você considera adequado o uso de IA para criar conteúdos sem informar que foram gerados por uma máquina?

- Sim, em qualquer contexto

- Sim, em situações específicas
- Não, deve haver transparência
- Não, o conteúdo deve sempre ser identificado como gerado por IA
- Não me sinto apto(a) a responder

7. Como você enxerga o uso de IA na criação de conteúdos potencialmente enganosos (como desinformação ou *deepfakes*)?

- É um problema sério e urgente
- É preocupante, mas pode ser controlado
- É um risco baixo se houver fiscalização
- Não vejo como um problema relevante
- Não me sinto apto(a) a responder

8. Você acredita que a IA pode reproduzir ou reforçar preconceitos existentes na sociedade ao gerar conteúdos?

- Sim, com frequência
- Sim, mas pode ser evitado com boas práticas
- Não, se os dados forem bem tratados
- Não, a IA é neutra por natureza
- Não me sinto apto(a) a responder

9. Na sua opinião, quem deve ser responsabilizado por eventuais danos causados por conteúdos gerados por IA?

- Quem utilizou a ferramenta
- A empresa ou desenvolvedora da IA
- Ambos, conforme o caso
- Nenhum, pois a IA não tem intenção própria
- Não me sinto apto(a) a responder

10. Você acredita que a IA pode afetar a criatividade humana no processo de criação digital?

- Sim, pode limitar a criatividade

- Sim, mas depende de como é utilizada
- Não, pode até estimular novas formas de criar
- Não, a criatividade humana é sempre predominante
- Não me sinto apto(a) a responder

11. Você considera as decisões tomadas por sistemas de IA confiáveis em sua área de atuação?

- Sim, totalmente
- Sim, mas precisam de supervisão humana
- Não, oferecem riscos importantes
- Não, a IA não deve tomar decisões sozinha
- Não me sinto apto(a) a responder

12. Você acha que o avanço da IA na produção digital deve continuar sendo estimulado?

- Sim, sem restrições
- Sim, com acompanhamento e normas claras
- Não, apenas com controle rigoroso
- Não, o avanço deve ser limitado
- Não me sinto apto(a) a responder

13. Imagine que você tem acesso a uma IA altamente avançada, capaz de responder a qualquer pergunta, mas que apresenta uma margem de erro conhecida. Qual seria o limite de erro aceitável para que você confiasse nas respostas?

- Até 60% de erro
- Até 40% de erro
- Até 25% de erro
- Até 10% de erro
- Não me sinto apto(a) a responder

APÊNDICE B - Entrevista

Parte I – Identificação e Contextualização

1. Poderia nos contar um pouco sobre sua trajetória profissional e sua relação com a Inteligência Artificial?
2. Como você avalia o papel da IA na sociedade atual?

Parte II – Ética e Regulação

3. Na sua visão, quais são os principais desafios éticos no uso de LLMs na produção de conteúdos digitais?
4. Você acredita que há necessidade de regulação específica para o uso desses modelos? Por quê?
5. Quem deve ser responsável por eventuais danos causados pelo uso inadequado de IA — o desenvolvedor, o usuário ou outro agente?

Parte III – Autoria, Transparência e Direitos Autorais

6. Como você vê a questão da autoria quando o conteúdo é gerado por IA? Deve haver atribuição específica?
7. A transparência no uso de IA é essencial? Em quais contextos isso é mais relevante?
8. Como a IA afeta os direitos autorais e a propriedade intelectual?

Parte IV – Impacto Social e Profissional

9. Em sua área de atuação, você acredita que a IA substituirá ou apenas complementarará o trabalho humano?
10. Como a IA pode influenciar a criatividade e a originalidade na produção de conteúdos?

Parte V – Considerações Finais

11. Qual sua opinião sobre o futuro da IA generativa na produção digital?
12. Que recomendações você daria para promover um uso mais ético e consciente dessa tecnologia?