



Faculdade de Tecnologia de Americana “Ministro Ralph Biasi”
Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas

**AMANDA ERRERA DE QUEIROZ
KAYKY DOS SANTOS MEDRADO
LARISSA GARCIA NUNES DE OLIVEIRA**

ANÁLISE DE FERRAMENTAS DE VERIFICAÇÃO DE IA
Uma aplicação em textos jornalísticos

Americana, SP

2025

**Amanda Errera de Queiroz
Kayky dos Santos Medrado
Larissa Garcia Nunes de Oliveira.**

ANÁLISE DE FERRAMENTAS DE VERIFICAÇÃO DE IA Uma aplicação em textos jornalísticos

Trabalho de Conclusão de Curso desenvolvido em cumprimento à exigência curricular do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas na área de concentração em Inteligência Artificial Aplicada.

Orientador: Prof. Mestre Dr.Rafael Rodrigo Martinati.

Este trabalho corresponde à versão base da monografia, contendo: introdução, referencial teórico e referências, do Trabalho de Conclusão de Curso apresentado por Amanda Errera de Queiroz, Kayky dos Santos Medrado e Larissa Garcia Nunes de Oliveira, orientado pelo Prof. Mestre Rafael Rodrigo Martinati.

**Americana, SP
2025**

**FICHA CATALOGRÁFICA – Biblioteca Fatec Americana
Ministro Ralph Biasi- CEETEPS Dados Internacionais de
Catalogação-na-fonte**

QUEIROZ, Amanda

Análise de ferramentas de verificação de IA: uma aplicação em textos jornalísticos . / Amanda Queiroz, Kayky Medrado, Larissa Oliveira – Americana, 2025.

130f.

Monografia (Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas) - - Faculdade de Tecnologia de Americana Ministro Ralph Biasi – Centro Estadual de Educação Tecnológica Paula Souza

Orientador: Prof. Ms. Rafael Rodrigo Martinati

1. Comunicação de massa – ética 2. Inteligência artificial 3. Jornalismo. I. QUEIROZ, Amanda , II. MEDRADO, Kayky, III. OLIVEIRA, Larissa IV. MARTINATI, Rafael Rodrigo V. Centro Estadual de Educação Tecnológica Paula Souza – Faculdade de Tecnologia de Americana Ministro Ralph Biasi

CDU: 659.3:17
007.52
659.3

Elaborada pelo autor por meio de sistema automático gerador de ficha catalográfica da Fatec de Americana Ministro Ralph Biasi.

Amanda Ferreira de Queiroz
Larissa Garcia Nunes de Oliveira
Kayky dos Santos Medrado

Análise de Ferramentas de Classificação de IA

Trabalho de graduação apresentado como exigência parcial para obtenção do título de Tecnólogo em Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas pelo Centro Paula Souza – FATEC Faculdade de Tecnologia de Americana Ministro Ralph Biasi.
Área de concentração: Análise e Desenvolvimento de Sistemas.

Americana, 8 de novembro de 2025.

Banca Examinadora:

Rafael Rodrigo Martinati
Mestre
Fatec Americana "Ministro Ralph Biasi"


Rogerio Nunes de Freitas
Mestre
Fatec Americana "Ministro Ralph Biasi"


Rodrigo Brito Battilana
Mestre
Fatec Americana "Ministro Ralph Biasi"

AGRADECIMENTOS

Agradecemos, em primeiro lugar, ao Professor Rafael Rodrigo Martinati, nosso orientador, pela dedicação, paciência e pelas valiosas contribuições que foram fundamentais para o desenvolvimento deste trabalho. Sua orientação e incentivo constante foram essenciais para que pudéssemos concluir esta pesquisa com confiança, aprendizado e uma perspectiva mais ampla sobre nosso papel como profissionais .

Às nossas famílias, pelo apoio incondicional, pela compreensão diante das ausências e pelo incentivo ao longo de toda a trajetória acadêmica. Sem o suporte emocional e a confiança de vocês, esta conquista não seria possível.

Por fim, agradecemos profundamente a nós mesmos, os três autores deste trabalho. Enfrentamos juntos prazos apertados, noites em claro, momentos de incerteza e ansiedade dos resultados. Mas também compartilhamos risadas, aprendizados e conquistas. A paciência, o carinho e o respeito que cultivamos entre nós transformaram essa jornada em algo muito maior do que um simples projeto acadêmico: foi uma experiência de crescimento, parceria e amizade. Cada um de nós trouxe força quando o outro precisou, e foi essa união que nos trouxe até aqui.

RESUMO

O presente trabalho investiga a confiabilidade de ferramentas de detecção de conteúdo gerado por Inteligência Artificial no contexto de textos jornalísticos. Com a crescente utilização de tecnologias de IA generativa, especialmente modelos de linguagem como ChatGPT, enfrentam-se desafios crescentes para verificar a autoria da produção textual jornalística. Este estudo tem como objetivo avaliar e comparar o desempenho de três ferramentas comerciais de detecção (ZeroGPT, GPTZero e Originality.AI), analisando sua precisão, confiabilidade e adequação para identificar textos de diferentes origens. A pesquisa adota metodologia quantitativa e comparativa, estruturada em quatro etapas principais: seleção de material textual diversificado, criação dos textos híbridos e artificiais, submissão às ferramentas de detecção e análise estatística dos resultados. O material compreende 12 textos distribuídos em três categorias: 100% humanos, 100% gerados por IA e híbridos. Para cada texto, foram registradas, a partir da submissão às três plataformas, as métricas percentuais de classificação. Os dados foram organizados em tabelas analíticas individuais e consolidados em médias por categoria e por ferramenta, possibilitando análise comparativa sistemática. O resultado obtido foi que as ferramentas de detecção não atendem ao percentual de acerto que elas propagam.

Palavras Chave: Inteligência Artificial; Detectores de IA; Textos Jornalísticos.

ABSTRACT

This study investigates the reliability of Artificial Intelligence-generated content detection tools in the context of journalistic texts. With the growing use of generative AI technologies, especially language models like ChatGPT, there are increasing challenges to verify the authorship of journalistic textual production. This study aims to evaluate and compare the performance of three commercial detection tools (ZeroGPT, GPTZero, and Originality.AI), analyzing their accuracy, reliability, and suitability for identifying texts from different origins. The research adopts a quantitative and comparative methodology, structured in four main stages: selection of diversified textual material, creation of hybrid and artificial texts, submission to detection tools, and statistical analysis of the results. The material comprises 12 texts distributed into three categories: 100% human, 100% AI-generated, and hybrid. For each text, the percentage classification metrics were recorded after submission to the three platforms. The data were organized into individual analytical tables and consolidated into averages per category and per tool, enabling systematic comparative analysis. The obtained result was that the detection tools do not achieve the accuracy percentages they propagate.

Keywords: *Artificial Intelligence; AI Detectors; Journalistic Texts.*

LISTA DE FIGURAS

Figura 1 - Diagrama de processamento de modelo de ML.....	20
Figura 2 - Processamento de Linguagem Natural e suas aplicações.....	22
Figura 3 - Hierarquia e Relações entre IA, ML, PLN e LLM.....	28
Figura 4 - Desempenho do GPT-5 em Diferentes Benchmarks (2025).....	37
Figura 5 - A Evolução dos Sistemas de Escrita.....	39
Figura 6 - Os Três Pilares da Escrita.....	40
Figura 7 - Tipos de Produção Textual na Era Digital.....	43
Figura 8 - Fórmulas de Equação linear utilizadas.....	70

LISTA DE TABELAS

Tabela 1 - Níveis de conhecimento linguístico.....	24
Tabela 2 - Definições de LLMS.....	28
Tabela 3 – Tempo estimado para atingir 1 milhão de usuários.....	32
Tabela 4 – Usuários ativos semanais do ChatGPT.....	34
Tabela 5 - Timeline evolutiva do GPT.....	35
Tabela 6 - Configuração do Delineamento Experimental.....	58
Tabela 7 - Tipologias de Notícia.....	59
Tabela 8 - Elementos que compõem do prompt.....	60
Tabela 9 - Elementos que compõem do prompt.....	61
Tabela 10 - Formato de preenchimento de prompt.....	62
Tabela 11 - Formato de preenchimento de prompt.....	64
Tabela 12 - Instrução Específica de Aprimoramento.....	65
Tabela 13 - Resumo dos Textos Componentes da Amostra para Análise.....	67
Tabela 14 - Tabela de termos.....	69
Tabela 15 - Análise Texto 100% Humano 1 - H1/Crítica.....	72
Tabela 16 - Análise Texto 100% Humano 2 - H2/Insólita.....	72
Tabela 17 - Análise Texto 100% Humano 3 - H3/Factual.....	73
Tabela 18 - Análise Texto 100% Humano 4 - H4/Dados.....	73
Tabela 19 - Análise Texto 100% IA 1 - IA1/Crítica.....	74
Tabela 20 - Análise Texto 100% IA 2 - IA2/Insólita.....	74
Tabela 21 - Análise Texto 100% IA 3 - IA3/Factual.....	75
Tabela 22 - Análise Texto 100% IA 4 - IA4/Dados.....	75
Tabela 23 - Análise Texto Híbrido 1 - HYB1/Crítica.....	76
Tabela 24 - Análise Texto Híbrido 2 - HYB2/Insólita.....	77
Tabela 25 - Análise Texto Híbrido 3 - HYB3/Factual.....	77
Tabela 26 - Análise Texto Híbrido 4 - HYB4/Dados.....	78

Tabela 27 - Média porcentagem dos textos 100 % humanos - H1,H2,H3 E H4.....	79
Tabela 28 - Média porcentagem dos textos 100 % IA - IA1,IA2,IA3 E IA4.....	79
Tabela 29 - Média porcentagem dos híbridos - HYB1,HYB2,HYB3 E HYB4.....	80
Tabela 30 - Análise por Ferramenta: ZeroGPT.....	80
Tabela 31 - Análise por Ferramenta: GPTZero.....	82
Tabela 32 - Análise por Ferramenta: Originality.....	83
Tabela 33 - Comparativo entre as Ferramentas de detecção.....	84
Tabela 34 - Comparativo dos resultados gerais.....	84

LISTA DE ABREVIATURAS E SIGLAS

IA	Inteligência Artificial
ML	Machine Learning
LLMs	Large Language Models
PLN	Processamento de Linguagem Natural
COPE	Committee on Publication Ethics
VP	Verdadeiro positivo
VN	Verdadeiro negativo
FP	Falso positivo
FN	Falso negativo

SUMÁRIO

LISTA DE FIGURAS.....	8
LISTA DE TABELAS.....	9
LISTA DE ABREVIATURAS E SIGLAS.....	11
1 . INTRODUÇÃO.....	14
2 . FUNDAMENTAÇÃO TEÓRICA.....	17
2.1 Machine learning.....	17
2.1.1 Modelos de Machine Learning : IA Generativa e IA Discriminativa.....	18
2.2 O Processamento de Linguagem Natural (PLN).....	21
2.2.1 Mecanismos de Atenção e tokenização.....	23
2.2.2 A Arquitetura Transformer.....	25
2.3 Grandes Modelos de Linguagem (LLMs).....	27
2.3.1 ChatGPT.....	32
2.3.1.1 GPT-5: Inovação e Limitações na Nova Geração.....	36
2.4 Produção Textual.....	38
2.4.1 A Escrita Antigamente.....	41
2.4.2 A Escrita Atualmente.....	42
2.4.3 Textos Híbridos.....	45
2.4.3.1 Textos Híbridos e Gerados por IA.....	45
2.4.3.2 Texto Híbrido ou Texto Humano Revisado.....	47
2.4.4 Limitações Técnicas.....	47
2.4.5 Desafios Técnicos e Jurídicos.....	48
2.5 O que é autoria?.....	49
2.5.1 Como era comprovado a autoria no passado.....	51
2.5.2 Como é comprovado atualmente.....	52
2.5.3 Quais são os limites da Autoria das IA e seu Uso Ético.....	53
2.5.4 Ferramentas de Detecção de Textos Gerados por IA.....	53

2.5.4.1 Ferramenta de Detecção de IA em textos (GPTZero).....	54
2.5.4.2 Ferramenta de Detecção de IA em textos (ZeroGPT).....	54
2.5.4.3 Ferramenta de Detecção de IA em textos (Originality.AI).....	55
2.5.5 Discussão crítica.....	55
3. METODOLOGIA.....	56
3.1 Tipo de Pesquisa e Abordagem.....	56
3.2 Objeto de Estudo e Ferramentas Empregadas.....	57
3.3 Delineamento Experimental e Geração de Dados.....	57
3.3.1 Critérios de Seleção e Geração dos Textos.....	58
3.3.1.1 Textos Gerados por Humanos (100% Humanos).....	59
3.3.1.2 Textos Gerados integralmente por Inteligência Artificial.....	59
3.3.1.3 Textos Híbridos.....	63
3.3.2 Resumo da Amostra Textual.....	67
3.3.3 Ambiente de Teste Controlado e Delimitação Temporal.....	68
3.4 Coleta e Análise de Dados.....	69
3.5 Limitações da Pesquisa.....	71
4. ANÁLISE DOS RESULTADOS.....	72
5. CONSIDERAÇÕES FINAIS.....	86
6. REFERÊNCIAS BIBLIOGRÁFICAS.....	88
APÊNDICE A – PROMPT UTILIZADO PARA GERAÇÃO DE TEXTO COM IA.....	94
APÊNDICE B – PROMPT UTILIZADO PARA GERAÇÃO DE TEXTO HÍBRIDO.....	95
APÊNDICE C – IA1 / Crítica.....	96
APÊNDICE D – IA2 / Insólita.....	98
APÊNDICE E – IA3 / Factual.....	100
APÊNDICE F – IA4 / Dados.....	102
APÊNDICE G – HYB 1 / Crítica.....	104
APÊNDICE H – HYB 2 / Insólita.....	106

APÊNDICE I – HYB 3 / Factual.....	109
APÊNDICE J – HYB 4 / Dados.....	111
ANEXO A – H1 / Crítica.....	115
ANEXO B – H2 / Isólita.....	117
ANEXO C – H3 / Factual.....	121
ANEXO D – H4 / Dados.....	123

1 . INTRODUÇÃO

A revolução digital das últimas décadas transformou radicalmente a forma como a informação é produzida, disseminada e consumida. No centro dessa transformação, a Inteligência Artificial emergiu como uma das tecnologias mais inovadoras do século XXI, redesenhando os limites entre a criação humana e a geração automatizada de conteúdo. Particularmente no campo da produção textual, modelos de linguagem de grande escala, como o ChatGPT, lançado pela OpenAI em novembro de 2022, democratizaram o acesso a ferramentas capazes de gerar textos coesos, gramaticalmente corretos e contextualmente relevantes em questão de segundos.

Esse avanço tecnológico, embora promissor em diversos aspectos, trouxe consigo desafios inéditos para múltiplos setores da sociedade. No jornalismo, área que historicamente se fundamenta na credibilidade, verificação factual e autoria reconhecível, a disseminação de textos gerados por IA levanta questões críticas sobre autenticidade, ética profissional e confiabilidade da informação. Instituições educacionais enfrentam dilemas similares ao avaliar trabalhos acadêmicos, enquanto o mercado editorial questiona os limites da originalidade e da propriedade intelectual em um cenário onde a distinção entre produção humana e artificial se torna cada vez mais difícil.

Diante desse contexto, surge uma demanda crescente por mecanismos confiáveis de identificação de conteúdo gerado por Inteligência Artificial. Diversas empresas de tecnologia desenvolveram ferramentas comerciais de detecção que prometem distinguir textos humanos de textos artificiais com altas taxas de precisão, oferecendo-se como soluções para instituições educacionais, veículos de comunicação e organizações que necessitam verificar a autoria de produções textuais. Essas plataformas, como ZeroGPT, GPTZero e Originality.AI, utilizam algoritmos de aprendizado de máquina treinados para identificar padrões linguísticos, estruturas sintáticas e o estilo de escrita potencialmente associadas à geração artificial.

Contudo, a eficácia real dessas ferramentas permanece objeto de debate acadêmico e profissional. A complexidade se intensifica quando consideramos textos

híbridos, aqueles produzidos colaborativamente entre humanos e IA, cenário cada vez mais comum na prática contemporânea de produção textual. Adicionalmente, a diversidade de gêneros textuais, estilos de escrita e contextos de produção pode influenciar significativamente a capacidade de detecção dessas ferramentas, levantando questionamentos sobre sua confiabilidade.

No contexto brasileiro, onde a discussão sobre verificadores de IA ainda é introdutória, torna-se fundamental produzir dados que sirvam de base para decisões institucionais e práticas profissionais. Diante desse cenário, emerge a seguinte questão de pesquisa: qual o nível de confiabilidade das ferramentas comerciais de detecção de conteúdo gerado por IA na identificação correta de textos jornalísticos de diferentes origens (humanos, artificiais e híbridos) em língua portuguesa?

Este trabalho tem como objetivo avaliar comparativamente a confiabilidade e a precisão de três ferramentas comerciais de detecção de conteúdo gerado por Inteligência Artificial (ZeroGPT, GPTZero e Originality.AI) na identificação de textos jornalísticos de diferentes origens.

Para alcançar esse propósito central, o estudo busca analisar o desempenho individual de cada ferramenta na classificação de textos 100% humanos, 100% gerados por IA e híbridos, observando como cada plataforma se comporta diante das três categorias de autoria. Além disso, pretende-se comparar as taxas de precisão, falsos positivos e falsos negativos das três ferramentas em diferentes tipos textuais jornalísticos, identificando padrões de desempenho e limitações específicas.

O trabalho também visa verificar a capacidade das ferramentas em identificar textos híbridos, produzidos colaborativamente entre humanos e IA, considerando que essa categoria representa um desafio particular para os algoritmos de detecção. Ademais, busca-se contrastar os resultados obtidos com as especificações técnicas e promessas comerciais declaradas pelas próprias ferramentas, avaliando a distância entre o desempenho anunciado e o desempenho real.

Justifica-se a relevância desta pesquisa pela necessidade de produzir dados sólidos que orientem o uso consciente e crítico dessas tecnologias no contexto brasileiro. Em um cenário onde instituições educacionais, veículos jornalísticos e organizações diversas recorrem crescentemente a verificadores automatizados, torna-se imperativo compreender suas reais capacidades e limitações. A confiança excessiva em ferramentas imperfeitas pode resultar em acusações injustas, desqualificação de trabalhos legítimos ou, inversamente, em falha na identificação de conteúdo efetivamente artificial.

Este trabalho está organizado em cinco capítulos. O primeiro capítulo apresenta esta introdução, contextualizando o tema, definindo o problema de pesquisa, os objetivos e a justificativa do estudo. O segundo capítulo desenvolve o referencial teórico, abordando conceitos fundamentais sobre Inteligência Artificial, produção textual, autoria, mecanismos de detecção de conteúdo artificial e suas aplicações no contexto jornalístico. O terceiro capítulo descreve detalhadamente a metodologia empregada, incluindo os critérios de seleção do corpus textual, os procedimentos de criação dos textos híbridos e articiais, detalhando a elaboração dos prompts utilizados e a forma como foram preenchidos, as etapas de submissão às ferramentas e os métodos de análise estatística dos dados. O quarto capítulo apresenta os resultados obtidos, organizando-os em tabelas analíticas individuais por texto, seguidos por análises consolidadas do desempenho médio geral das ferramentas e uma análise detalhada da performance de cada uma delas, oferecendo diferentes perspectivas sobre os resultados, com descrições objetivas de seus achados. Por fim, o quinto capítulo traz as considerações finais, retomando os objetivos, sintetizando os principais achados, discutindo suas implicações práticas e sugerindo caminhos para pesquisas futuras.

2 . FUNDAMENTAÇÃO TEÓRICA

2.1 Machine learning

O aprendizado de máquina (em inglês *machine learning*), é um dos campos da Inteligência Artificial que mais transformou o cenário computacional. Através dele, os sistemas passaram a aprender de maneira autônoma com os dados recebidos, ajustando parâmetros a cada novo processamento.

Diferente da programação tradicional, que depende unicamente de instruções fixas, o aprendizado de máquina (ML) consegue “aprender” e reconhecer padrões, adaptar a resposta com base em experiências anteriores e assim se tornar cada vez mais independentes, tornando-se capaz de criar conteúdos novos que seguem estilos e formatos semelhantes aos dados utilizados. Isso se tornou cada vez mais notório, por exemplo, na produção textual.

Para que essa autonomia e capacidade de criação de conteúdo sejam plenamente desenvolvidas, a qualidade e diversidade dos dados de treinamento tornam-se essenciais, como afirma Garcia (2020, p.16) “no aprendizado de máquina, os dados desempenham um papel fundamental: quanto mais dados (confiáveis) disponíveis para treinar o algoritmo, melhor será o modelo gerado por ele”.

No caso da escrita, isso significa que bases textuais amplas e consistentes são essenciais não apenas para garantir correção gramatical, mas também para assegurar coerência temática, adequação do estilo de escrita e flexibilidade diante de diferentes contextos comunicativos. Ademais, a qualidade dos dados é decisiva, já que vieses presentes no material de treinamento, como preconceitos sociais ou representações desproporcionais de determinados grupos, podem ser assimilados e reproduzidos pelo modelo, resultando em produções tendenciosas ou injustas.

Nesse ínterim, é necessário destacar que uma das limitações , como reforçam FORD (2018 *apud* KAUFMAN e SANTAELLA, 2020, p. 3), é :

[...] esses sistemas carecem da essência da inteligência humana: capacidade de compreender o significado; igualmente não possuem senso intuitivo, capacidade de formar conceitos abstratos e de fazer analogias e generalizações, não têm a capacidade de compreender o funcionamento do mundo a partir da observação (conceitos como tridimensionalidade, movimentação e permanência dos objetos, gravidade, inércia e rigidez, dentre outros) (LECUN citado por FORD, 2018).

Deste modo, embora possam gerar textos fluidos e coerentes na aparência, os sistemas de IA operam por meio da manipulação estatística, sem alcançar compreensão real do conteúdo. Portanto, seu desempenho depende diretamente da qualidade e da consistência das bases de dados, consequentemente podem acabar produzindo mensagens discriminatórias ou de conteúdo ilícito, caso identifiquem tais padrões como válidos em sua base de dados, sem que haja a devida atribuição de responsabilidade, dada a sua incapacidade de entendimento. Isso constitui uma limitação fundamental e os diferencia essencialmente dos textos produzidos por seres humanos, que escrevem a partir de sua bagagem vivencial e assumem responsabilidade ética e legal pelo que comunicam.

Portanto, mesmo que os sistemas de IA consigam realizar tarefas antes restritas à inteligência humana, continuam presos à reprodução de padrões já aprendidos, sem atingir o tipo de compreensão profunda que caracteriza o raciocínio humano.

2.1.1 Modelos de Machine Learning : IA Generativa e IA Discriminativa

Os modelos de *machine learning* (ML) podem ser organizados em diferentes categorias metodológicas. Neste trabalho, serão abordadas duas abordagens específicas: os modelos discriminativos e generativos. Em outras palavras, não estamos lidando com algoritmos ou categorias independentes de IA, mas de estratégias distintas de processamento de dados.

O modo como esses modelos lidam com dados pode ser ilustrado pela analogia apresentada por LeCun, citado por Ford (2018, p.124, tradução própria):

Você mostra a imagem de um cavalo e, se ele não disser 'cavalo', você diz que está errado e aqui está a resposta que ele deveria ter dito. Então, usando o algoritmo de retropropagação, ele ajusta todos os pesos de todas as conexões na rede para que, da próxima vez que você mostrar a mesma imagem de cavalo, a saída seja mais próxima da desejada, e você continua fazendo isso para milhares de imagens.¹

Essa explicação evidencia que, em ambas as abordagens, os modelos aprendem a partir de exemplos e de ajustes sucessivos nas conexões internas. O que muda, entretanto, é o tipo de saída produzida e a finalidade do aprendizado.

Os modelos discriminativos funcionam como classificadores especializados cujo trabalho é aprender a separar diferentes tipos de dados, desenvolvendo durante o treinamento a habilidade de identificar características que distinguem uma categoria da outra. Pense em um sistema que diferencia gatos de cachorros: ele aprende que orelhas pontudas, bigodes e certos formatos de focinho tendem a indicar "gato", enquanto orelhas caídas, focinhos alongados e determinadas posturas sugerem "cachorro". Quando apresentado a uma nova imagem, o modelo simplesmente consulta esse "manual de diferenças" que construiu e emite seu veredito. A qualidade desse processo depende diretamente do treinamento recebido, modelos bem alimentados com dados diversos não apenas captam diferenças evidentes, como tamanho ou cor, mas também detectam detalhes que passariam despercebidos ao olhar humano desatento.

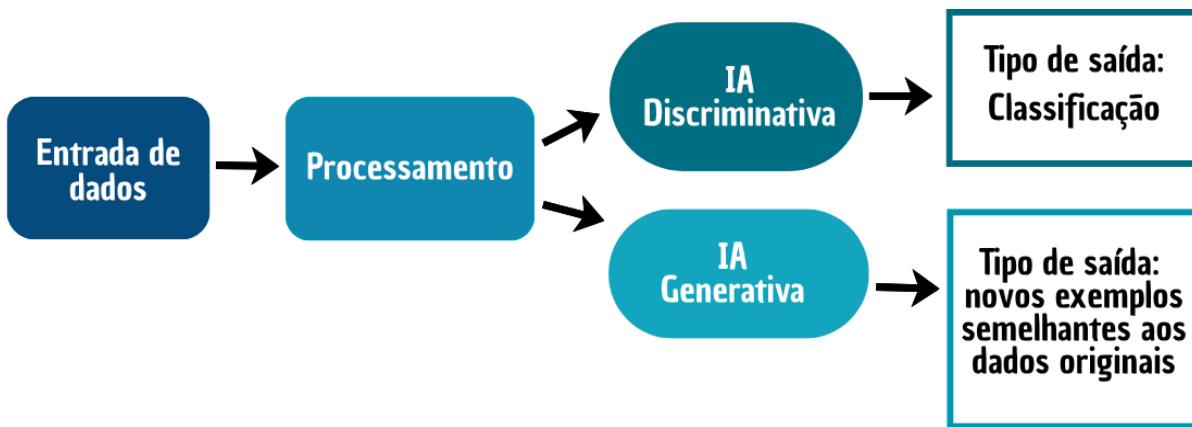
Em contraste, os modelos gerativos não se restringem a classificar; eles procuram entender como os dados estão organizados e distribuídos para então criar novos exemplos similares aos originais. Conforme exemplificado por Goodfellow et al. (2020, p.139, nossa tradução), “O objetivo de um modelo gerativo é estudar um conjunto de exemplos de treinamento e aprender a distribuição de probabilidade que os gerou.”².

¹ No original : "You show it the image of a horse, and if it doesn't say 'horse,' you tell it that it's wrong and here is the answer that it should have said. Then by using the backpropagation algorithm, it adjusts all the weights of all the connections in the network so that next time you show the same image of a horse, the output would be closer to the one you want, and you keep doing this for thousands of images."

² No original : "The goal of a generative model is to study a collection of training examples and learn the probability distribution that generated them"

A “IA generativa, estas são capazes de sintetizar conteúdo, comumente textos e imagens, com base em comandos dos usuários” (Dwivedi et al., 2023 *apud* TRINDADE; OLIVEIRA, 2024, p.9). Desse modo, a IA generativa se destaca pela capacidade de criar novos dados que mantêm as características dos dados de treinamento. É importante destacar, contudo, que “ IA Generativa não conseguem gerar novas ideias” (TRINDADE; OLIVEIRA, 2024, p.14), mas sim recombinar e produzir informações já existentes em seu conjunto de dados de treinamento.

Figura 1 - Diagrama de processamento de modelo de ML



Fonte : Elaborado pelos autores.

A diferença fundamental entre os dois tipos de modelos reside no resultado que produzem, como observado na Figura 1. Enquanto os modelos discriminativos traçam fronteiras para separar categorias, como distinguir gatos de cachorros em uma imagem, os modelos generativos aprendem a estrutura dos dados para produzir algo novo a partir do que conhecem, como gerar uma imagem inédita de gato baseada em milhares de exemplos vistos anteriormente. Esta distinção determina suas aplicações: discriminativos classificam, generativos criam.

2.2 O Processamento de Linguagem Natural (PLN)

Entre os muitos desafios da Inteligência Artificial, compreender a linguagem humana talvez seja o elemento-chave mais complexo de superar. Afinal, a comunicação humana vai muito além de juntar palavras, ela carrega cultura, emoção e contexto.

Expressões como “a cobra vai fumar” exemplificam essa complexidade, apesar de conterem termos como “cobra” e “fumar”, seu significado real nada tem a ver com répteis ou tabaco; qualquer falante nativo do português brasileiro entende que indica uma confusão iminente. De forma semelhante, um simples “Kabum!” dispensa tradução literal para transmitir uma explosão. Uma máquina pode conhecer o significado isolado de cada palavra, mas captar essas camadas interpretativas exige muito mais do que vocabulário: é preciso expô-la a uma variedade imensa de exemplos e contextos para que comece a perceber nuances que os humanos assimilam naturalmente desde a infância.

Para enfrentar esse desafio, surge o Processamento de Linguagem Natural. “O processamento de linguagem natural emprega técnicas computacionais com o objetivo de aprender, compreender e produzir conteúdo da linguagem humana.” (HIRSCHBERG; MANNING, 2015, p. 261, nossa tradução)³. Nesse sentido, o objetivo primordial, portanto, é capacitar os sistemas a realmente compreenderem, interpretarem e reagirem à fala de forma intuitiva e natural.

A relevância do PLN se acentua drasticamente frente ao volume crescente de dados textuais produzidos diariamente. Schopow, Osterhoff e Baur (2023, n.p., nossa tradução) destacam que o PLN " surgiu como uma ferramenta poderosa [...], permitindo o processamento e a análise de grandes quantidades de dados textuais não

³ No original : "Natural language processing employs computational techniques for the purpose of learning, understanding, and producing human language content."

estruturados em vários domínios".⁴ Essa capacidade de transformar dados desorganizados em informações úteis consolidou o PLN como elemento indispensável na era digital.

Atualmente, os sistemas de PLN são capazes de compreender, interpretar e gerar linguagem humana de forma abrangente e sofisticada. Isso abrange “diferentes ramos os quais esta área pode se dividir devido à sua extensa gama de possibilidades, como recuperação de informação, tradução e interpretação de textos, análise de sentimentos, etc.” (VIEIRA; LOPES, 2010 *apud* PAES, 2025, p. 5). Como demonstrado na Figura 2, o PLN viabiliza múltiplas aplicações práticas: recuperação de informações em grandes bases de dados, análise de sentimentos para identificar opiniões, extração de informações para capturar dados específicos, tradução automática entre idiomas, e sistemas de perguntas e respostas baseados em conhecimento textual.

Figura 2 - Processamento de Linguagem Natural e suas aplicações



Fonte : Elaborado pelos autores

⁴ No original : "Natural Language Processing (NLP) has emerged as a powerful tool in recent years, enabling the processing and analysis of vast amounts of unstructured textual data in various domains."

Esses avanços permitiram tanto o progresso de sistemas capazes de gerar texto com alto grau de sofisticação, ocasionando o surgimento dos Modelos de Linguagem de Grande Escala (LLMs), os quais serão discutidos na seção 2.3.

2.2.1 Mecanismos de Atenção e tokenização

No contexto do processamento de linguagem natural, os modelos computacionais enfrentam uma dificuldade marcante: a limitada capacidade de diferenciar e priorizar informações cruciais em suas sequências de dados. Tal restrição decorria do processamento tradicional, que conferia idêntica relevância a cada elemento de entrada, inviabilizando que os modelos se concentrassem nos detalhes verdadeiramente essenciais para um determinado contexto.

Essa lacuna foi superada com a introdução dos mecanismos de atenção. Conforme Stryker e Bergmann (s.d., s.p.) definem que "Um mecanismo de atenção é uma técnica de aprendizado de máquina que direciona modelos de deep learning para priorizar (ou atender) às partes mais relevantes dos dados de input". Apresentados inicialmente por Bahdanau et al. (2014), esses mecanismos capacitaram os modelos a focar seletivamente nas porções mais pertinentes da entrada para cada tarefa específica.

O conceito por trás da atenção tem uma inspiração biológica natural. Stryker e Bergmann (s.d., s.p.) explicam que "Como o nome sugere, os mecanismos de atenção são inspirados na capacidade dos seres humanos (e de outros animais) de se concentrar seletivamente em detalhes relevantes e ignorar os menos importantes no momento". Essa capacidade se manifesta quando, ao buscar uma informação específica, a atenção humana se volta naturalmente aos trechos mais pertinentes, ignorando as informações secundárias. De modo análogo, a atenção capacita os modelos de PLN a "observar" as sequências de entrada com diferentes graus de relevância.

Para que o mecanismo de atenção possa atuar com precisão, a linguagem computacional demanda uma preparação crucial: a tokenização. É neste ponto que os textos são desmembrados em suas unidades mais elementares, os tokens⁵. Tais tokens são os elementos básicos que os algoritmos de PLN interpretam e processam. É por meio deles que “os modelos processam a linguagem tentando entender a estrutura gramatical (sintaxe) e o significado (semântica). O objetivo é produzir uma linguagem com a sintaxe e a semântica corretas que sejam relevantes para a entrada” (SHASTRI, 2025, s.p.). Para atingir essa compreensão da linguagem, os modelos consideram os diferentes níveis de conhecimento linguístico, conforme detalhado na Tabela 1.

Tabela 1 - Níveis de conhecimento linguístico

Nível de Conhecimento	Definição
Fonético e fonológico	Do relacionamento das palavras com os sons que produzem.
Morfológico	Da construção das palavras a partir de unidades de significado primitivas e de como classificá-las em categorias morfológicas.
Sintático	Do relacionamento das palavras entre si, cada uma assumindo seu papel estrutural nas frases, e de como as frases podem ser partes de outras, constituindo sentenças.
Semântico	Do relacionamento das palavras com seus significados e de como eles são combinados para formar os significados das sentenças.
Pragmático	Do uso de frases e sentenças em diferentes contextos, afetando o significado.

Fonte: (GONZALEZ; LIMA, 2003, adaptado pelo Autores).

⁵ Tokens são elementos que podem ser palavras inteiras, partes delas ou até caracteres, dependendo da abordagem de segmentação escolhida.

Tecnicamente, a atenção funciona ao calcular as pontuações de relevância (scores) para cada elemento da sequência de entrada. Essas pontuações são dinamicamente atribuídas com base no estado atual do processamento. Após a normalização, transformam-se em pesos que quantificam a "atenção" a ser direcionada a cada elemento. O resultado é uma representação ponderada que enfatiza as informações mais pertinentes para o contexto, permitindo ao modelo processar ao mesmo tempo os diferentes níveis linguísticos detalhados na Tabela 1.

Para ilustrar, considere a tradução da frase "O gato subiu no telhado" para o inglês. No momento em que o modelo necessita gerar a palavra "roof" (telhado), o mecanismo de atenção concentra-se prioritariamente na palavra "telhado" da frase original. A capacidade de direcionar o foco ao significado específico (nível semântico) é fundamental. Entretanto, a atenção também avalia como as palavras se conectam na frase e qual é o sentido completo que a tradução pretende transmitir, pois a compreensão geral da sentença depende de como todas as suas palavras funcionam em conjunto.

Esse avanço teve impacto transformador em diversas aplicações de PLN, especialmente em tradução automática, onde os modelos passaram a produzir traduções significativamente mais precisas e contextualmente adequadas, estabelecendo as bases para desenvolvimentos ainda mais revolucionários.

2.2.2 A Arquitetura Transformer

O campo do Processamento de Linguagem Natural (PLN), impulsionado pelos mecanismos de atenção, ascendeu a um novo patamar com a emergência da arquitetura Transformer. Lançada em 2017 por pesquisadores da Google no artigo "Attention is All You Need", esta arquitetura revolucionou a área ao conceber um design que se fundamenta exclusivamente em atenção.

A inovação primordial do Transformer concentra-se no conceito de *self-attention* (autoatenção), que expandiu e aprimorou a ideia dos mecanismos de atenção já conhecidos. Diferente da atenção convencional, que primordialmente estabelecia relações entre a saída e a entrada, a *self-attention* permite que cada elemento de uma sequência busque ativamente e considere a relevância de todos os outros elementos dentro daquela mesma sequência. Essa interação abrangente possibilita ao modelo estabelecer dependências internas, independentemente da proximidade posicional dos *tokens*. Esta é uma habilidade crucial, pois, como descrevem Galassi, Lippi e Torroni (2021 *apud* PAES, 2025, p. 11):

Este mecanismo é responsável por conectar diferentes partes da sentença de entrada para criar uma representação da sentença livre dos limites posicionais das palavras no texto, permitindo que o modelo seja capaz de aprender palavras e características de uma sentença independentemente de sua posição relativa às outras partes da sentença

Na prática, isso permite que uma palavra, mesmo no início de uma frase, estabeleça conexões diretas e robustas com outra no final, sem que a distância física na sequência comprometa a força dessa relação. Essa abordagem representa uma ruptura radical com as arquiteturas predecessoras, que processavam as sequências de forma linear, perdendo progressivamente informações conforme a distância aumentava.

As vantagens dessa abordagem são notáveis. Primeiramente, sua estrutura permite o processamento paralelo, o que significa que o Transformer consegue processar múltiplas partes do texto simultaneamente. Isso agiliza drasticamente o treinamento, otimizando o uso do hardware moderno. Em segundo lugar, o Transformer supera um desafio persistente em modelos de linguagem anteriores: a dificuldade em reter informações cruciais ao longo de sequências extensas. Sua engenharia permite que cada palavra interaja diretamente com qualquer outra no texto, independentemente de sua posição. Com isso, o modelo consegue preservar um entendimento contextual rico e consistente do começo ao fim, sem que informações importantes se percam ou enfraqueçam com a distância.

Para materializar essa capacidade, a arquitetura Transformer se estrutura em dois módulos complementares: o *encoder* e o *decoder*. O encoder atua como um leitor perspicaz, encarregado de analisar e compreender profundamente a sequência de entrada, por exemplo, uma frase a ser traduzida. Ele processa todas as palavras, identificando suas relações e o contexto global. Em seguida, o decoder assume a tarefa de gerar a sequência de saída, utilizando a compreensão detalhada fornecida pelo *encoder*. Ambos os módulos incorporam as camadas de *self-attention*, que são a chave para essa capacidade de entender e produzir linguagem de forma tão eficiente.

Para melhor compreensão de seu funcionamento, considere a frase "O estudante que estava na biblioteca emprestou o livro". Em arquiteturas recorrentes, ao processar a palavra "emprestou", o modelo poderia ter atenuado a relevância de informações sobre "estudante". No Transformer, contudo, "emprestou" estabelece conexões diretas com "estudante" e "biblioteca" de forma simultânea, preservando integralmente as relações contextuais.

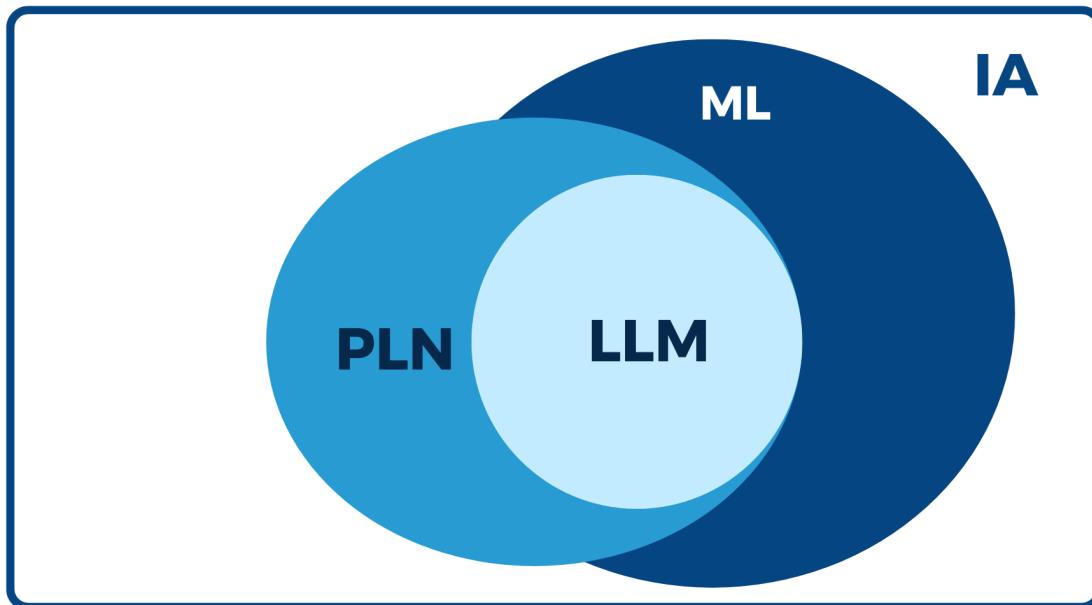
O impacto do Transformer foi notável e de longo alcance. Além de aprimorar significativamente o desempenho em tarefas convencionais de PLN, a arquitetura demonstrou uma capacidade ímpar de escalabilidade: quanto maiores os modelos e a quantidade de dados de treinamento, mais robusto se torna seu desempenho. Essa característica de escalabilidade provou ser fundamental para a ascensão dos Modelos de Linguagem de Grande Escala (LLMs) subsequentes.

2.3 Grandes Modelos de Linguagem (LLMs)

Dando continuidade aos avanços fundamentais dos Mecanismos de Atenção e da arquitetura Transformer, discutidos nas seções anteriores (2.2.1 e 2.2.2), os Modelos Grandes de Linguagem (LLMs) surgem como um ponto culminante no Processamento de Linguagem Natural (PLN). Tais sistemas, ao otimizarem radicalmente o processamento de relações contextuais e superarem limitações de dependência posicional, inauguraram uma nova era na interação homem-máquina,

demonstrando capacidades sem precedentes na compreensão e geração de linguagem humana.

Figura 3 - Hierarquia e Relações entre IA, ML, PLN e LLM



Fonte : Elaborado pelos autores

Fundamentalmente, um LLM é um sistema computacional sofisticado, um tipo de modelo de *deep learning* que aprendeu padrões linguísticos a partir de um volume massivo de textos. Dada a complexidade dessa tecnologia, diferentes autores apresentam definições complementares, conforme apresentado na Tabela 2 .

Tabela 2 - Definições de LLMS

Autor	Definição
Minaee et al. (2024, s.p., tradução própria)	Os grandes modelos de linguagem [...] contêm dezenas a centenas de bilhões de parâmetros, que são pré-treinados em dados de texto massivos [...] exibem habilidades de compreensão e geração de linguagem mais fortes e, mais

	importante, habilidades emergentes que não estão presentes em modelos de linguagem de menor escala. ⁶
Idan A. Blank (2023, p.987, tradução própria)	os LLMs são os modelo computacional mais bem implementado para processamento de linguagem natural: eles são os primeiros sistemas que lidam com linguagem natural em escala. ⁷
Kasneci et al., (2023 <i>apud</i> ALMARIE et al. 2023, p.1, tradução própria)	Um sistema de inteligência artificial (IA) estreita treinado em uma quantidade massiva de dados de texto para interpretar a linguagem natural e gerar respostas semelhantes às humanas para prompts ou perguntas baseadas em texto. ⁸
Manning (2022 <i>apud</i> MEYER et al. 2023, tradução própria)	Um tipo de modelo de deep learning que surgiu por volta de 2018, treinado em quantidades massivas de dados de texto publicamente disponíveis para gerar respostas semelhantes às humanas em conversas. ⁹

Fonte: Elaborado pelos autores com base em Minaee et al. (2024), Blank (2023), Kasneci et al. (2023 *apud* ALMARIE et al., 2023) e Manning (2022 *apud* MEYER et al., 2023).

A Tabela 2 demonstra, em síntese, que um LLM constitui um sofisticado modelo de deep learning, treinado em vastos volumes de dados e com bilhões de parâmetros. Essa abordagem permite-lhe não apenas interpretar e compreender a linguagem natural, mas também gerar respostas e textos coerentes, semelhantes aos produzidos por humanos, em conversas e outras interações.

Nesse contexto, a arquitetura sofisticada que compõe os LLMs permitiu ao modelo alcançar uma versatilidade notável. Além da geração de texto, esses sistemas expandem suas aplicações para um amplo conjunto de funcionalidades. Como destacam Matarazzo e Torloni (2025, p.4, nossa tradução) : "Seu desempenho

⁶ No original : "Large language models [...] contain tens to hundreds of billions of parameters, which are pre-trained on massive text data [...] exhibit stronger language understanding and generation abilities, and more importantly, emergent abilities that are not present in smaller-scale language models."

⁷ No original : "LLMs are the best implemented computational model for natural language processing: they are the first systems that handle natural language at scale."

⁸ No original : "A large language model (LLM) is a narrow artificial intelligence (AI) system that has been trained on a massive amount of text data to interpret natural language and generate human-like responses to text-based prompts or questions"

⁹ No original : "ChatGPT is powered by a LLM, a type of deep learning model that emerged around 2018. These models are trained on massive amounts of publicly available text data, such as books, articles, and webpages, to generate human-like responses in conversations."

impressionante abrange uma variedade de tarefas, incluindo geração de texto, resposta a perguntas, tradução de idiomas e sumarização, demonstrando seu potencial para enfrentar desafios linguísticos complexos.". ¹⁰

Ademais, ao processar volumes massivos de informação e aprender as complexas relações dentro da linguagem, eles exibem capacidades que simulam um raciocínio mais sofisticado. Matarazzo e Torlone (2025, p.4, nossa tradução) complementam: "Esses modelos [...] revolucionaram a forma como abordamos tarefas de processamento de linguagem natural, alcançando níveis de compreensão, aprendizagem e geração que antes eram considerados inatingíveis."¹¹ Tais manifestações enfatizam a complexidade do treinamento e a riqueza dos dados que modelam esses sistemas.

Contudo, é crucial entender que a "inteligência" dos LLMs reside na manipulação estatística de padrões, característica inerente aos sistemas de Machine Learning, conforme detalhado na seção 2.1., esses modelos operam identificando padrões em seus vastos dados de treinamento, sem, no entanto, possuir compreensão intrínseca ou raciocínio humano. Tal distinção é vital, pois, como Meyer et al.(2023, p.2, nossa tradução) ressaltam:

"É preciso lembrar que não é um humano: não é treinado para responder de uma forma que reflete com precisão suas próprias capacidades e limitações, mas sim, é treinado para construir enunciados textuais modelados a partir de outros semelhantes encontrados em seu texto de treinamento dado o prompt."¹²

Essa natureza estatística evidencia-se notavelmente na produção textual que aborda sentimentos ou opiniões. Um LLM pode gerar um poema melancólico, um discurso apaixonado ou uma crítica mordaz, mas o faz replicando padrões linguísticos

¹⁰ No original : "Their impressive performance spans a variety of tasks, including text generation, question answering, language translation, and summarization, showcasing their potential in tackling intricate language challenges."

¹¹ No original : "These models [...] have revolutionized how we approach natural language processing tasks, achieving comprehension, learning, and generation levels that were once considered unattainable."

¹² No original : "One needs to remember that it is not a human: it's not trained to respond in a way that accurately reflects its own capabilities and limitations, but rather, it is trained to construct textual utterances modeled after similar ones found in its training text given the prompt.".

associados a esses estados emocionais ou posições ideológicas em seu treinamento. Ele não "sente" a melancolia, a paixão ou a indignação, nem "acredita" na crítica que formula. A capacidade de "escrever sobre sentimentos sem senti-los" ou de "gerar uma opinião sem ser capaz de realmente ter uma opinião" é uma prova da sofisticação de sua modelagem linguística, mas também um lembrete constante de sua natureza algorítmica.

Em outras palavras, isso significa que um texto pode soar bem escrito, mas não necessariamente ser verdadeiro. Embora um LLM possa gerar textos gramaticalmente impecáveis e coerentes, como Meyer et al. (2023, p.2, nossa tradução) apontam, "uma resposta gerada por um modelo de linguagem grande pode ser formatada corretamente, mas não necessariamente ser factual"¹³. Isso reforça a necessidade crítica de avaliação e verificação humana das informações produzidas por esses sistemas, especialmente em contextos onde a precisão é indispensável.

Plataformas como ChatGPT da OpenAI, Gemini do Google e Claude da Anthropic representam a materialização prática dos LLM. Para ilustrar, pode-se comparar um LLM a uma biblioteca vasta e bem organizada, junto a um bibliotecário que memorizou cada detalhe de seu acervo. O ChatGPT, nesse cenário, atua como um assistente especializado que atende no balcão de informações. Ao receber uma solicitação, o assistente (ChatGPT) consulta a "biblioteca" (o LLM) para construir uma resposta coerente e útil, sem, contudo, ser o criador original do conhecimento.

Nesse contexto de inovação, o ChatGPT se destaca como um dos exemplos mais proeminentes, tornando-se uma interface conversacional amplamente reconhecida, impulsionada pelos Modelos Grandes de Linguagem.

¹³ No original : "an answer generated by a large language model can be formatted correctly but not necessarily be factual".

2.3.1 ChatGPT

No meio de uma grande mudança dos Modelos Grandes de Linguagem (LLMs) reside a arquitetura GPT (Generative Pre-trained Transformers). Treinada extensivamente em grandes volumes de texto, essa classe de modelos aprende a prever a próxima palavra, gerando linguagem humana notavelmente coerente e contextualmente relevante. Essa capacidade é impulsionada pela arquitetura Transformer, que, com seu processamento paralelo e mecanismos de atenção sofisticados, confere aos GPTs uma capacidade ímpar de compreender e criar textos. Essa inovação abriu o caminho para o ChatGPT da OpenAI. Sua interface intuitiva e fácil de usar popularizou a inteligência artificial avançada, integrando-a ao dia a dia de milhões. Este cenário não só transformou profundamente a dinâmica da criação e análise textual, mas também acendeu um alerta para a crescente necessidade de ferramentas robustas que identifiquem a autoria de textos.

A chegada do ChatGPT redefiniu a produção textual assistida por IA, marcando um divisor de águas na tecnologia. Sua capacidade de diálogo natural e resposta imediata, garantiu uma aceitação e acessibilidade sem precedentes, resultando em um aumento na produção de conteúdo automatizado e na diversificação dos tipos de texto gerados por IA. A Tabela 3 ilustra com clareza essa trajetória impressionante: o ChatGPT alcançou 1 milhão de usuários em meros cinco dias. Um contraste marcante se estabelece com plataformas consolidadas, como Instagram (dois meses), Facebook (dez meses) e Twitter (dois anos). Essa agilidade na adoção não apenas revelou uma demanda latente por IAs interativas, mas também consolidou a presença do ChatGPT na escrita diária. Com isso, surgem questões complexas sobre a autoria e a identificação de textos produzidos por máquinas.

Tabela 3 – Tempo estimado para atingir 1 milhão de usuários

Plataforma	Tempo para atingir 1 milhão de usuários
Tópicos do Instagram	1 hora
ChatGPT	5 dias

Instagram	2 meses
Spotify	5 meses
Dropbox	7 meses
Facebook	10 meses
Quatro Quadrados	13 meses
Twitter	2 anos
Airbnb	2,5 anos
Kickstarter	2,5 anos
Netflix	3,5 anos

Fonte: Shubham Singh (2025) adaptado pelos autores

A facilidade de uso do ChatGPT o tornam uma ferramenta valiosa, criando uma crescente dependência para diversas atividades. Essa integração da IA não se limita ao uso pessoal, tornando-se um recurso fundamental no ambiente de trabalho. A própria OpenAI (2025) detalha esse uso: "O uso do ChatGPT pelos consumidores é, em grande parte, voltado para a realização de tarefas cotidianas. Três quartos das conversas se concentram em orientações práticas, busca de informações e escrita — sendo a escrita a tarefa de trabalho mais comum, enquanto codificação e autoexpressão continuam sendo atividades de nicho." Esses dados revelam como a IA se tornou uma ferramenta essencial para a produção textual contemporânea.

Essa profunda inserção nas rotinas pessoais e profissionais impulsionou um crescimento notável em sua base de usuários ativos. A Tabela 4 ilustra essa expansão: de 50 milhões de usuários ativos semanais em janeiro de 2023, o número saltou para expressivos 800 milhões em abril de 2025. Essa adoção massiva evidencia o impacto transformador do ChatGPT, enquanto a interação constante dessa vasta comunidade gera um volume inédito de dados e feedback, que realimenta e aprimora continuamente os modelos GPT subjacentes.

Tabela 4 – Usuários ativos semanais do ChatGPT

Mês	Usuários ativos semanais do ChatGPT
Abril de 2025	800 milhões
Fevereiro de 2025	400 milhões
Dezembro de 2024	300 milhões
Outubro de 2024	250 milhões
Agosto de 2024	200 milhões
Agosto de 2023	100 milhões
Janeiro de 2023	50 milhões

Fonte: Shubham Singh (2025)

A performance linguística dos modelos GPT e sua notável habilidade em gerar textos cada vez mais sofisticados derivam diretamente do aumento maciço de seus parâmetros, que são as variáveis internas que o modelo assimila durante o treinamento. Quanto mais parâmetros, mais complexos se tornam os padrões e nuances da linguagem que o modelo consegue discernir e replicar, tornando a tarefa de distinguir entre a produção humana e a artificial um desafio crescente para os identificadores de IA.

Essa correlação entre parâmetros e capacidade linguística torna-se evidente na evolução histórica dos modelos GPT. A Tabela 5 detalha essa progressão: o GPT-1 (2018), com 117 milhões de parâmetros, focava na previsão da próxima palavra. O GPT-4 (2023) alcançou 170 trilhões de parâmetros. Este salto resultou em um aprimoramento linguístico substancial, permitindo a geração e edição de textos criativos e técnicos com proficiência, e até a assimilação de estilos de escrita específicos.

Tabela 5 - Timeline evolutiva do GPT

Modelo e ano	Contagem de Parâmetros	Habilidades
GPT-1 (2018)	117 milhões de parâmetros	O GPT-1 demonstrou o poder da aprendizagem não supervisionada em tarefas de compreensão de linguagem, usando livros como dados de treinamento para prever a próxima palavra em uma frase. (MARR, 2023, s.p., tradução própria). ¹⁴
GPT-2 (2019)	1,5 bilhão de parâmetros	Ele apresentou uma melhoria drástica nas capacidades de geração de texto e produziu textos coerentes com vários parágrafos. (MARR, 2023, s.p., tradução própria). ¹⁵
GPT-3 (2020)	175 bilhões de parâmetros	Seus recursos avançados de geração de texto levaram ao seu amplo uso em diversas aplicações, desde a elaboração de e-mails e a escrita de artigos até a criação de poesias e até mesmo a geração de código de programação (MARR, 2023, s.p., tradução própria). ¹⁶
GPT-4 (2023)	170 trilhões de parâmetros	Ele pode gerar, editar e interagir com usuários em tarefas de redação criativa e técnica, o que significa que pode ser usado para compor canções, criar roteiros e até aprender o estilo de escrita de um usuário. (OPENAI, s.d., s.p.)
GPT-5 (2025)	Informação indisponível	O GPT-5 é mais inteligente em todos os parâmetros, oferecendo respostas mais úteis em matemática, ciências, finanças, direito e muito mais. É como ter uma equipe de especialistas à disposição para tudo que você quer saber. (OPENAI, 2025, s.p.)

Elaborado pelos autores com base em Marr (2023) e OpenAI (s.d.; 2025).

¹⁴ No original : “GPT-1 demonstrated the power of unsupervised learning in language understanding tasks, using books as training data to predict the next word in a sentence.”

¹⁵ No original : “ It showcased a dramatic improvement in text generation capabilities and produced coherent, multi-paragraph text.”

¹⁶ No original : “ Its advanced text-generation capabilities led to widespread use in various applications, from drafting emails and writing articles to creating poetry and even generating programming code

A busca incessante por aprimoramento da arquitetura GPT culminou no GPT-5, a geração mais recente e avançada desses modelos. Ele promete um novo patamar de desempenho e funcionalidade para a inteligência conversacional, aproximando-se ainda mais da produção textual da linguagem humana. Esse avanço, por sua vez, adiciona outra camada de complexidade aos sistemas de identificação de IA.

2.3.1.1 GPT-5: Inovação e Limitações na Nova Geração

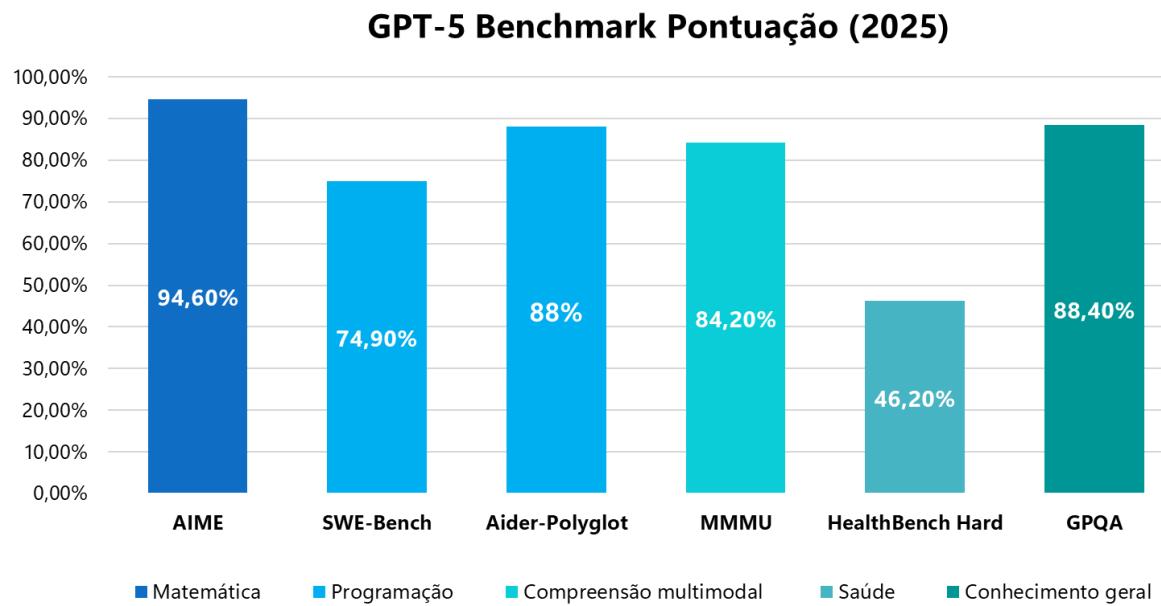
O lançamento do ChatGPT-5, em 7 de agosto de 2025, é a versão mais avançada de modelos de linguagem (LLMs) da OpenAI. Embora a contagem exata de parâmetros não tenha sido divulgada, é inegável que o GPT-5 incorpora um aumento considerável em sua complexidade e capacidade. Com o GPT-4 já operando em 170 trilhões de parâmetros, pode-se inferir que o GPT-5 eleva ainda mais o patamar de poder computacional e inteligência artificial, resultando em textos mais precisos e versáteis.

Para compreender melhor esse aprimoramento, é importante analisar como a OpenAI posiciona este novo modelo. A empresa descreve o GPT-5 como um "sistema unificado", incorporando um "modelo inteligente e eficiente" capaz de responder a um leque ampliado de questões e demonstrar "reflexão para problemas mais difíceis". Seus aprimoramentos incluem um "direcionador em tempo real", que refina interações com base na complexidade e intenção explícita do usuário. Segundo a OpenAI (2025), o GPT-5 é 'como ter uma equipe de especialistas à disposição', demonstrando seu avanço em múltiplas áreas.

Traduzindo essas promessas em resultados concretos, os avanços do GPT-5 em termos de desempenho são substanciais e mensuráveis. O modelo demonstra "melhor desempenho que os modelos anteriores em relação a benchmarks" e promete "responder mais perguntas mais rapidamente" (OpenAI, 2025, s.p.). Esses benchmarks são testes padronizados para avaliar o desempenho de sistemas de IA e incluem avaliações em áreas como compreensão de leitura, raciocínio matemático, codificação

e conhecimento geral. Como pode ser visto na Figura 4 , essas avaliações demonstram o escopo abrangente das capacidades do GPT-5 em diferentes domínios de conhecimento.

Figura 4 - Desempenho do GPT-5 em Diferentes Benchmarks (2025)



Elaborado pelos autores (2025), com dados do GPT-5 (OpenAI, 2025).

Na prática, isso significa que o GPT-5 consegue processar consultas complexas com maior velocidade, oferece respostas mais precisas em tarefas especializadas, demonstra melhor compreensão contextual em conversas longas e mantém maior consistência lógica ao longo de suas respostas. Essa melhoria quantificável indica não apenas maior eficiência computacional, mas também uma capacidade aprimorada de produzir conteúdo que se aproxima ainda mais da qualidade e coerência da escrita humana.

Apesar desses avanços impressionantes, uma característica crucial desta nova versão é sua maior confiabilidade, impactando diretamente os desafios de identificação. Especificamente em relação às alucinações, geração de informações incorretas ou sem sentido, mas apresentadas de forma convincente, a OpenAI (2025,

s.p.) afirma: "O ChatGPT também alucina, embora no GPT-5 isso aconteça com frequência bem menor, principalmente quando a reflexão está ativada. As alucinações continuam sendo um desafio fundamental para todos os grandes modelos de linguagem, mas estamos trabalhando com afinco para reduzi-las". Embora esta redução represente um avanço vital na busca por maior veracidade nas respostas, a própria OpenAI reconhece que o modelo não está completamente isento dessas falhas, indicando que o avanço técnico, por mais significativo que seja, não elimina completamente os desafios inerentes aos modelos de linguagem.

Esses avanços tecnológicos do GPT-5 geram implicações significativas para o cenário da produção textual contemporânea. A crescente sofisticação do modelo, aliada à redução das alucinações e ao aumento da veracidade, cria um paradoxo interessante: quanto mais eficiente e confiável se torna a IA generativa, mais desafiadora se torna a tarefa de distinguir textos gerados artificialmente daqueles produzidos por humanos. À medida que o texto gerado se aproxima da qualidade da escrita humana, os sistemas de identificação de IA enfrentam a tarefa cada vez mais complexa de discernir nuances sutis.

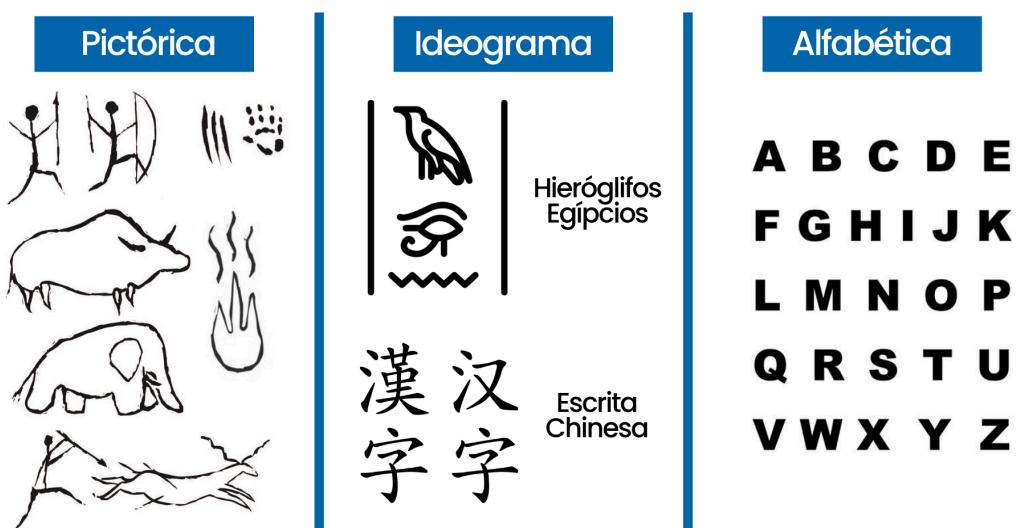
Diante desse contexto, surgem questões fundamentais sobre autoria, originalidade e autenticidade na era digital, temas que se tornam cada vez mais relevantes à medida que a produção textual assistida por IA se populariza. As implicações dessas transformações para os processos de criação de conteúdo, bem como os desafios emergentes para a detecção e verificação de autoria, serão explorados nas seções 2.4 e 2.5, que abordarão respectivamente como essas tecnologias estão redefinindo a produção textual e os métodos desenvolvidos para identificar conteúdo gerado por inteligência artificial, questões centrais para a discussão sobre autoria na era da IA generativa.

2.4 Produção Textual

A escrita, como registro da linguagem, é um dos principais meios pelos quais o ser humano deixou sua mensagem para aqueles que viriam depois dele. Os primeiros

registros eram ainda os desenhos rupestres das cavernas, na época neolítica, e desde então a escrita passou por diversas fases. Segundo Sampson (1996 *apud* ANDRADE, 2001), a escrita surgiu após a revolução neolítica e pode ser compreendida em três fases principais: pictórica, ideográfica e alfabética (Figura 5). A primeira refere-se a desenhos que representavam diretamente objetos; a segunda, aos ideogramas que expressavam ideias, como os hieróglifos egípcios e a escrita chinesa; e a terceira, ao uso de letras que passaram a ter função fonética, consolidando o sistema alfabético.

Figura 5 - A Evolução dos Sistemas de Escrita



Fonte : Elaborado pelos autores .

A partir de então, a escrita firmou-se como a ferramenta primordial para a expressão, registro e transmissão do conhecimento. Isso permitiu que ideias, experiências e saberes fossem compartilhados e preservados, impulsionando a evolução das sociedades e servindo de alicerce para novas descobertas, pensamentos e teorias, culminando em significativas conquistas literárias, científicas e tecnológicas. A invenção da escrita representou um avanço essencial para o desenvolvimento humano, como salienta Andrade (2001, p. 4):

O que se pode dizer com total convicção é que a invenção da escrita foi um grande avanço para o desenvolvimento da humanidade, pois ela representa nossas ideias que podem ficar registradas por muitos e muitos anos, diferentemente da fala, que, se não for gravada, brevemente se desvanece.

Nesse contexto, a escrita permanece como o meio pelo qual o ser humano se organiza como indivíduo e como sociedade. Bortolanza e Costa (2016, p.4-5) complementam:

Pela linguagem, o homem influencia externamente o outro e internamente a si mesmo. Esse processo se realiza externamente, por meio da comunicação oral e escrita, pelo intercâmbio de informações, e internamente, pela posterior propriedade de converter-se em meio de orientação das operações intelectuais.

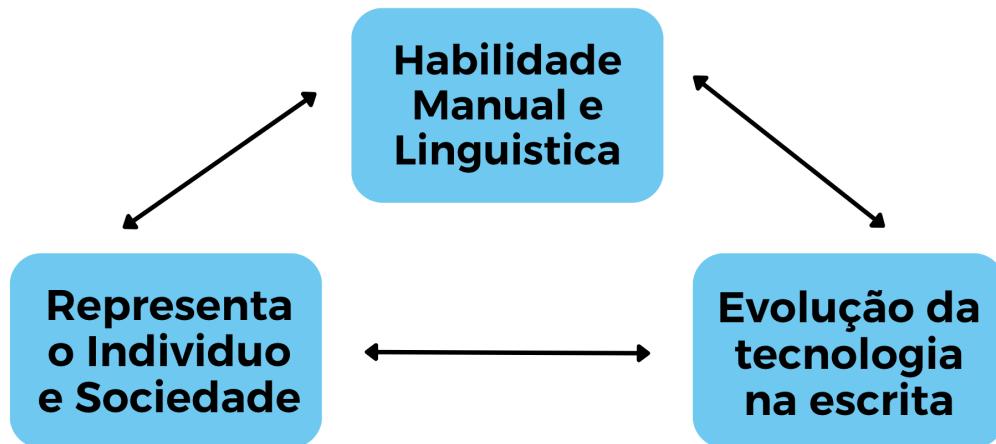
Diante dessa importância, é fundamental que os membros de uma sociedade tenham domínio da escrita, de modo a compreender sua cultura e participar ativamente dela, tanto para usufruir dos benefícios sociais, políticos e econômicos quanto para contribuir com o desenvolvimento do seu meio social. Segundo Marques (2001 *apud* NOGARO; PORTO; PORTO, 2019), a escrita não é apenas registro de pensamentos prévios, mas um processo que estimula o raciocínio lógico, a reflexão e a construção do autoconhecimento e da identidade do autor.

Com o avanço tecnológico atual e seu impacto na escrita, torna-se ainda mais relevante incentivar e cultivar a qualidade da produção textual. Como ressaltam Gnerre (1998 *apud* NOGARO; PORTO; PORTO, 2019):

Refletir sobre a escrita torna-se quase que uma necessidade diante da difusão em nível mundial da escrita na educação básica e, ao mesmo tempo, da sua prática em moldes e formas tradicionais, que parecem tornar a atividade obsoleta diante do avanço das novas tecnologias que impulsionam novos relacionamentos do sujeito com a escrita.

Finalmente, podemos afirmar que a escrita é essencial para o desenvolvimento humano, tanto individual quanto coletivo. Sendo um dos pilares da sociedade, ela não apenas forma o público que a recebe, mas também se constrói a partir das habilidades humanas que a produzem. Assim, são as habilidades humanas que determinam o futuro da sociedade e da escrita, duas áreas que se fortalecem mutuamente.

Figura 6 - Os Três Pilares da Escrita



Fonte : Elaborado pelos autores .

2.4.1 A Escrita Antigamente

A produção textual passou por transformações profundas ao longo da história humana. A escrita era um processo inteiramente manual e solitário, onde o autor concebia, desenvolvia e refinava suas ideias através de múltiplos rascunhos manuscritos ou datilografados. Grandes obras literárias, científicas e filosóficas emergiam de processos extensos e meticulosos, envolvendo múltiplas etapas de concepção, desenvolvimento e refinamento.

Tradicionalmente, escrever era um processo lento e organizado em etapas, da ideia aos rascunhos e revisões. Como explicam Silva e Freesz (2017, p. 2):

Todo texto literário se constrói a partir de um processo criativo. Sabemos que uma obra se forma, não apenas quando sua publicação acontece, mas também por todo o percurso de seu autor até o produto final, o livro. Para que o texto se constitua, são necessários caminhos que passam por anotações, rabiscos, esboços, versões datilografadas ou digitadas.

Nesse contexto, o escritor dependia exclusivamente de suas habilidades linguísticas, conhecimento prévio e experiência pessoal para construir narrativas, argumentos e textos técnicos. As ferramentas disponíveis limitavam-se a instrumentos

básicos como pena, tinta, papel e, posteriormente, máquinas de escrever. Consequentemente, a revisão e correção demandavam reescrita completa de seções ou páginas inteiras, tornando o processo moroso e trabalhoso.

Ademais, a pesquisa requeria consultas físicas a bibliotecas, enciclopédias e outras fontes impressas, o que naturalmente restringia o acesso à informação e prolongava significativamente o tempo necessário para produzir textos bem fundamentados. Finalmente, a obra era submetida ao escrutínio da comunidade intelectual relevante, seja através de publicação em periódicos, apresentação em círculos acadêmicos ou distribuição em redes sociais específicas.

Além do valor literário, esses registros do processo criativo possuem importância histórica e documental. Silva e Freesz (2017, p. 2) destacam que:

A reunião destes fragmentos, anteriores à publicação do livro, permite desvendar curiosidades a respeito da vida do autor, suas influências, suas fontes, suas sugestões, compreendendo o caminho da criação até a edição publicada. Ainda, por meio deles, estamos diante dos documentos da gênese e do estudo da memória, tanto do próprio texto em si, quanto de seu contexto: os costumes e hábitos referentes à época de sua criação.

Portanto, o processo de escrita tradicional refletia não apenas a essência e mensagem da obra, mas também a identidade de seu autor e o contexto histórico-cultural de sua época. Durante séculos, o desenvolvimento de textos manteve-se como um processo longo e extenuante, cenário que começou a se transformar gradualmente com o avanço da tecnologia.

2.4.2 A Escrita Atualmente

Nos dias de hoje, com a evolução tecnológica no estado em que estamos, diversas ferramentas foram desenvolvidas para ajudar na criação e desenvolvimento de textos. Antes mesmo de grandes marcos como a apresentação das IAs, com a globalização e a digitalização, já vinham sendo utilizadas pequenas ferramentas que auxiliavam um indivíduo a produzir um texto, como sites de brainstorms, dicionários de sinônimos e repositórios de matérias e artigos online.

Contudo, com a chegada das IAs, elas introduziram uma nova dimensão à escrita, onde algoritmos podem auxiliar, complementar ou até mesmo gerar conteúdo textual. Segundo Pimentel, Azevedo e Carvalho (2023, s.p.), "entendemos, assim, que iniciamos uma nova era em que a autoria se tornou híbrida humana/o-IA, em que as obras serão cocriadas, coproduzidas com a inteligência artificial criativa".

Esta transformação não apenas acelerou o processo de produção, mas também democratizou o acesso a ferramentas sofisticadas de escrita, permitindo que indivíduos com diferentes níveis de habilidade linguística produzam textos mais refinados e bem estruturados.

Nesse contexto, a produção textual atualmente pode ser dividida em três tipos principais que refletem a autoria e o nível de intervenção tecnológica: a humana pura , a inteiramente gerada por Inteligência Artificial (IA) e a híbrida. Esta categorização é essencial para entender as nuances e os desafios emergentes na validação de conteúdo na era digital.

Figura 7 - Tipos de Produção Textual na Era Digital



Fonte : Elaborado pelos autores .

2.4.2.1 Gênero Jornalístico: Notícia

Dentro do universo da produção textual, o gênero jornalístico, especialmente o formato notícia, merece destaque por suas características próprias e sua relevância social. A notícia tem como principal objetivo informar o público sobre acontecimentos recentes, sendo construída a partir de fatos verificáveis e apresentando uma estrutura definida. Diferente de textos opinativos ou literários, a notícia privilegia a objetividade, a clareza e a precisão, tornando-se uma ferramenta essencial para a circulação de informações na sociedade.

A linguagem jornalística é direta e concisa, buscando transmitir informações de forma acessível e comprehensível para diferentes públicos. Segundo Lustosa (1996 *apud* SILVA, 2008, p. 66), "a notícia é um produto simbólico destinado ao consumo da massa e, por isso, é feita para todo mundo a partir de uma técnica de produção capaz de ser absorvida por todos" . A objetividade e a linearidade do texto contribuem para que a notícia cumpra sua função social: permitir que os leitores compreendam os eventos e tomem decisões fundamentadas a partir das informações recebidas.

No contexto do uso de inteligência artificial na escrita, o gênero notícia apresenta vantagens e desafios específicos. Por ser estruturado e padronizado, ele facilita a avaliação de textos híbridos, permitindo observar como a IA se comporta na manutenção de coerência, fidelidade aos fatos e estilo jornalístico. Entretanto, a necessidade de precisão e verificação constante dos dados expostos ressalta a importância de supervisão humana, evidenciando que, mesmo com o auxílio de tecnologias, o controle editorial e a responsabilidade ética permanecem centrais (NOGARO; PORTO; PORTO, 2019). Assim, o estudo da produção de notícias híbridas

entre humanos e IA revela tanto as potencialidades quanto os limites da tecnologia na preservação da objetividade e do rigor informativo.

2.4.3 Textos Híbridos

Pela definição, híbrido significa: "Que possui progenitores cujos genótipos, composição genética, são diferentes; mestiço: indivíduo híbrido" (DICIO, 2025, s.p.). Logo, textos híbridos, no contexto do uso de IAs, são aqueles que envolvem interação dinâmica entre humano e máquina, onde o autor humano mantém controle editorial sobre o processo, direcionando a IA, avaliando suas sugestões e incorporando ou rejeitando suas contribuições algorítmicas.

Com isso, surgem novos questionamentos sobre a validação de textos híbridos, como identificá-los e determinar suas autorias. A Turnitin (2023) observa que as ferramentas de escrita por IA podem ser valiosas para apoiar a criatividade, mas tornam-se problemáticas quando o texto gerado é apresentado pelos alunos como se fosse de autoria própria, o que se aproxima da fraude acadêmica.

Para compreender a autoria de um texto que possui suspeita de uso de IA sem declaração dela, deveria-se saber o que é um texto feito inteiramente por IAs e como identificá-los. Em outra matéria, a Turnitin (2023, s.p.) complementa:

Precisamos ser cautelosos e saber que podemos identificar o uso dessas ferramentas quando elas não são aceitáveis. Quando tivermos esses guardiões em vigor, no entanto, será hora de voltarmos nossa atenção para as novas possibilidades que a tecnologia venha a apresentar.

2.4.3.1 Textos Híbridos e Gerados por IA

A distinção entre textos híbridos genuinamente colaborativos e textos inteiramente gerados por IA torna-se crucial para compreender como classificar um texto e, assim, entender sua autoria e valor.

Diferentemente dos textos híbridos, textos gerados inteiramente por IAs devem ser aqueles onde a intervenção humana se limita ao fornecimento de prompts iniciais e, possivelmente, à edição posterior mínima. São textos que apresentam claros padrões de escrita de IA e sem dados relevantes, originais ou inovadores.

Como já mencionado anteriormente, textos híbridos requerem uma participação ativa do autor humano, então a grande questão é qual o limite dessa participação necessária. Um texto que foi escrito por IA, mas supervisionado, guiado e editado por um coautor humano pode ser considerado um texto híbrido? Atualmente, não há uma explicação concreta, e a conclusão é que se faz necessário criar métricas para a identificação da origem e desenvolvimento do texto. De acordo com Jesuegraciliano (2024), é necessária uma contínua realização de debates e reflexões sobre o assunto e a implementação de normas referentes ao uso de IA no contexto acadêmico. E, de acordo com Pimentel, Azevedo e Carvalho (2023, s.p.):

Consideramos ser esse um (re)direcionamento importante que a educação deve assumir em tempo de Inteligência Artificial Criativa: superar a "educação bancária", que privilegia a memorização e repetição de conteúdos e de técnicas (FREIRE, 1970), e investir mais em atividades genuinamente autorais híbridas, críticas e criativas (CARVALHO; PIMENTEL, 2020).

Neste trabalho, considera-se como textos híbridos todos aqueles que sofreram influência, em algum grau, de um sistema de Inteligência Artificial durante seu processo de produção. Optou-se por essa definição ampla em razão da impossibilidade de estabelecer os níveis exatos de influência de IA e de contribuição humana em cada texto. A ausência de métricas precisas que permitam quantificar o grau de intervenção artificial torna inviável uma classificação mais granular, conduzindo à adoção de uma categorização binária: textos que receberam qualquer tipo de assistência de IA são classificados como híbridos, enquanto aqueles produzidos exclusivamente por humanos ou integralmente por IA constituem as demais categorias de análise.

2.4.3.2 Texto Híbrido ou Texto Humano Revisado

Um dos usos principais da IA deveria ser auxiliar na produção de textos de forma automatizada, aumentando a produtividade sem interferir de forma perceptível. Por isso, o uso da IA em revisões gramaticais e de coesão é um dos usos recomendados por profissionais e acadêmicos como um bom uso da ferramenta. O que se supõe limitado à não alteração de palavras, intenção, sentido e essência do texto original. Porém esse mesmo pequeno suporte pode fazer uma grande diferença no tom e direção do texto. Nesse contexto, deve-se levar em consideração que "como já bem estabelecido na literatura especializada, cada gênero do discurso apresenta convenções próprias, que moldam a organização da informação, a escolha lexical, a estrutura sintática e os recursos argumentativos empregados na construção dos enunciados linguísticos (Bakhtin, 1997)" (ANTONELLI, 2025, p. 2). E com isso, a IA deveria ser capaz de atender as especificações do tipo do texto que está sendo produzido. Porém nem sempre todos esses detalhes são levados em consideração na hora em que uma IA produz uma escrita. Atualmente as IAs ainda sofrem de falhas que comprometem a qualidade de um texto, e as principais serão descritas no tópico seguinte.

2.4.4 Limitações Técnicas

A IA, apesar de ser uma grande ferramenta generativa, ainda comete erros que podem ser gravemente prejudiciais no uso para a escrita. O que se faz questionar: seria uma IA realmente capaz de auxiliar ou mesmo produzir textos de qualidade equiparáveis aos humanos de forma consistente?

As limitações técnicas da inteligência artificial na produção textual vão além dos debates sobre autoria, envolvendo sobretudo a confiabilidade do que é gerado. Apesar dos avanços, ainda persistem falhas estruturais que comprometem a qualidade e exigem análise crítica para um uso responsável.

Entre essas falhas, destaca-se a tendência das IAs a reproduzirem padrões previsíveis, resultando em textos corretos, mas pouco originais e sem voz própria. Outro problema recorrente é a geração de informações falsas, as chamadas "alucinações", que podem incluir dados, citações e referências inexistentes apresentadas com aparente credibilidade, exigindo constante verificação.

A produção também depende fortemente da formulação dos prompts, já que pequenas variações nos comandos geram resultados muito diferentes em termos de qualidade e abordagem. Essa instabilidade se soma à ausência de compreensão semântica real: a IA apenas reconhece padrões estatísticos, sem entender de fato os conceitos, o que limita análises mais profundas e conexões conceituais amplas.

Quando confrontada com tarefas de raciocínio complexo, a tecnologia revela ainda mais suas fragilidades, como incoerências e conclusões inconsistentes, especialmente em textos acadêmicos que exigem rigor analítico. Além disso, seus limites refletem diretamente os dados de treinamento, que carregam vieses, lacunas e desatualizações, perpetuando problemas de precisão e representatividade.

2.4.5 Desafios Técnicos e Jurídicos

Atualmente, não existe no Brasil uma lei específica que regulamente de forma clara o uso de inteligências artificiais na produção textual. Como observam Marson e Cavalcante (2023, p.102), "há dificuldade em definir e proteger os direitos autorais de autores que utilizaram de IA em suas criações, já que se coloca em dúvida até que ponto a obra pode ser considerada humana, criativa e original." O que há são debates e projetos de lei em andamento, como o PL 2.338/2023, que busca estabelecer um marco legal para a IA no país, inspirado em regulações internacionais, como o AI Act da União Europeia, que já prevê regras de transparência e rotulagem de conteúdos gerados por máquinas. Além disso, setores da indústria criativa pressionam para que haja proteção aos direitos autorais, exigindo que criadores sejam informados e possam consentir quando suas obras forem utilizadas no treinamento dessas tecnologias.

Mesmo com essas tentativas de regulamentação, surgem desafios práticos de aplicação. Como destaca Caiado (2024, s.p.), "identificar a autoria e responsabilizar quem escreveu hoje em dia é muito difícil porque os atuais métodos que checam plágio usam de comparação de coisas que já foram apresentadas para eles. É muito difícil identificar plágio quando se usa inteligência artificial e o texto que você está usando é novo, nunca foi publicado." A exigência de informar o uso de IA em textos, por exemplo, depende quase sempre da honestidade do usuário, já que os métodos de detecção atuais apresentam limitações graves.

Ferramentas de identificação baseadas em padrões linguísticos cometem erros frequentes, classificando falsamente textos humanos como gerados por IA ou aceitando como humanos textos artificiais reformulados. As próprias empresas que desenvolvem ferramentas de detecção reconhecem que os programas não são totalmente confiáveis, pois frequentemente classificam textos humanos como gerados por IA e vice-versa (R7, 2024, s.p.). Soluções como marcas d'água digitais¹⁷, propostas por algumas empresas, podem ser facilmente contornadas por simples reescritas.

Assim, mesmo que seja possível identificar o uso de inteligência artificial em um texto, ainda não há critérios capazes de determinar sua autoria nem de avaliá-lo de forma justa quanto à sua contribuição para o conhecimento coletivo. Para avançar nessa discussão, é preciso compreender o que se entende por autoria, assunto da próxima seção.

2.5 O que é autoria?

Segundo a Lei nº 9.610, que rege o direito autoral no Brasil e está lei determina todos os direitos morais, intelectuais e patrimoniais dos autores sobre suas obras, conforme seu artigo 11: "Autor é a pessoa física criadora de obra literária, artística ou científica" (BRASIL, 1998). Enxergando o ponto de vista de Foucault (1969), o autor

¹⁷ A marca d'água digital é uma técnica utilizada para ocultar informações dentro de conteúdos digitais.

não é apenas o escritor de textos ou artigos, o autor é o responsável por aquela ideia que escreve. O autor carrega tanto os créditos da obra como as responsabilidades que estão atreladas a ela.

Ao direcionar o olhar para a autoria em textos e artigos científicos, a definição de autoria é definida por diversas instituições acadêmicas, tanto em nível nacional e nível internacional, Segundo o Committee on Publication Ethics (COPE) (2019, s.p.):

[...] O termo autoria pode se referir ao criador ou originador de uma ideia (por exemplo, o autor da teoria da relatividade) ou ao indivíduo ou indivíduos que desenvolvem e concretizam o produto que divulga obras intelectuais ou criativas. A autoria transmite privilégios, responsabilidades e direitos legais significativos;

Conforme as informações disponibilizadas pelo COPE (2019), As diretrizes de autoria mais conhecidas que são adotadas pelos periódicos biomédicos, Organizações editoriais, profissionais e diferentes academias, as recomendações do International Committee of Medical Journal Editors (ICMJE) a autoria tem que ser baseada nos quatro seguintes critérios: (1) Contribuições para concepção , aquisição, análise ou interpretação dos dados do trabalho. (2) Elaboração de versões preliminares ou revisões críticas importantes sobre o conteúdo descrito. (3) Aprovação final da versão a ser publicada; (4) Concordância em ser responsável por todos os aspectos do trabalho.

Em periódicos de instituições acadêmicas sazonalmente se destaca a importância da autoria e junto com ela o cuidado com o plágio, que não se limita à cópia do trabalho de outra pessoa, ele abrange o uso de qualquer produto fruto da inteligência de uma pessoa sem atribuir a devida autoria. “Portanto, parafrasear um texto, reformular métodos, gráficos e ideias sem acrescentar nada pessoal e sem atribuir devidamente a autoria é plágio” (Salvagno; Taccone; Gerli, 2023 *apud* TRINDADE; OLIVEIRA, 2024, p.10).

Em vista disso, diversas instituições acadêmicas procuraram estabelecer seus critérios sobre autoria, com o objetivo de regulamentar de forma correta e completa a identificação dos autores de artigos científicos, bem como seus respectivos créditos e

responsabilidades. E como essa regulamentação está sendo feita hoje em dia com a inserção de IA's como ferramentas nas produções textuais e como era feita essa regulamentação antigamente.

2.5.1 Como era comprovado a autoria no passado

A necessidade da autoria sempre esteve ligada a identificar, comprovar e responsabilizar o criador de uma obra intelectual. Antes da era digital, a comprovação da autoria se dava principalmente por meios físicos e sociais. Jornais, registros em cartório, testemunhos presenciais eram uma das formas de alegar a legitimidade de quem produzia a obra. Em instituições acadêmicas e científicas a legitimidade era feita pela publicação de periódicos impressos, que eram feitos por jornais e editoras que representavam a validação da autoria daquele texto ou obra (CHARTIER, 1998).

A comprovação de autenticidade nas produções acadêmicas estava sempre associada ao nome do autor e ao registro da instituição que foi feita a obra, seja em bibliotecas, universidades ou editoras. E sua assinatura era compreendida como marca de propriedade intelectual e reconhecimento de autenticidade (ECO, 1989).

Ao analisar a evolução das diretrizes da autoria científica se enxerga que a credibilidade e autenticidade dependia tanto da assinatura e nome do autor quanto da instituição responsável pela publicação.

Com isso, pode-se afirmar que antes da era digital, a comprovação de autoria se apoiava fortemente em registros físicos, assinaturas e as instituições garantem sua autoria, porém com o advento das tecnologias digitais e a inteligência artificial como é feita essa comprovação atualmente.

2.5.2 Como é comprovado atualmente

Com a chegada da era digital, foi notada a necessidade de comprovação de autoria e passaram a utilizar ferramentas que oferecem maior segurança e confiabilidade. Um dos sistemas utilizados para este fim atualmente é o Digital Object Identifier (DOI), desenvolvido pela CrossRef. O DOI é uma ferramenta que busca e identifica as publicações e artigos científicos de forma única e permanente, armazenando todas as obras pesquisadas nele e registrando elas em um código exclusivo, assim garantindo sua autenticidade e fácil citação em bases de dados acadêmicas (PENTZ, 2011, s.p.).

Outro recurso muito utilizado hoje em dia para fins autorais é o blockchain. Seu objetivo é criar e utilizar redes distribuídas e imutáveis, assim possibilitando o blockchain a armazenar de forma segura as informações relacionadas à autoria, impedindo alterações futuras e mantendo a proteção intelectual daquela obra armazenada (SWAN, 2015).

Além destes métodos, o uso de softwares para a detecção de plágio tem se popularizado em instituições de ensino e pesquisa. Ferramentas como Turnitin, CopySpider e GPTZero são utilizadas para verificar a autenticidade dos trabalhos acadêmicos, comparando textos contidos em um vasto banco de dados que contém diversas publicações, sites e documentos acadêmicos disponíveis online. Hexham (1992), antes da popularização da internet, Hexham já avisava da importância de conseguir identificar cópias não autorizadas em ambientes acadêmicos. Mais recentemente, Bretag (2019) fortaleceu esse debate ao relacionar o plágio com LLMs e a contratação de terceiros para realizar os trabalhos, assim aumentando a discussão sobre a necessidade de sistemas mais avançados para preservar a autenticidade acadêmica.

Atualmente, a comprovação da autoria está constantemente ligada à digitalização, tanto por meio de identificadores únicos e registros descentralizados, e com a chegada da IA houve um aumento crescente na dúvida de quais são os limites da autoria da IA e como é feito seu uso ético.

2.5.3 Quais são os limites da Autoria das IA e seu Uso Ético

Sob a visão jurídica, a legislação brasileira atualmente a lei que rege a autoria (Lei nº 9.610/1998) não contempla a possibilidade de autoria por IA, reconhecendo apenas pessoas físicas e jurídicas como sujeitos a cumprir a lei. Isso fortalece a ideia do uso ético de IA pois mesmo quando os textos são produzidos com auxílio de LLMs, a responsabilidade ética e legal é aplicada ao usuário que a utilizou para a produção textual.

Algumas instituições acadêmicas têm moldado e estabelecido novas diretrizes para lidar com essa questão. O Committee on Publication Ethics (COPE, 2023) recomenda que autores que utilizam IAs na produção ou coleta de dados devem ser transparentes ao publicar os materiais e métodos assim informando qual foi a ferramenta utilizada. O COPE também adverte que as IAs não podem atender aos requisitos de autoria pois não podem assumir a responsabilidade do que é criado, como entidades não legais a IA não pode gerenciar os direitos autorais ou licenças do que é feito por ela (COPE, 2023).

2.5.4 Ferramentas de Detecção de Textos Gerados por IA

Com a popularização do ChatGPT foram criadas as necessidades de descobrir se os conteúdos escritos eram ou não feitos por inteligência artificial e com isso vieram a criação de ferramentas específicas para detectar produções textuais geradas por LLMs, a maior parte da análise feita por esse detectores funcionam a partir de padrões linguísticos que normalmente são utilizados em texto artificiais.

De forma técnica, uma das principais métricas utilizadas é a “perplexidade”, que verifica a previsibilidade de um texto. Os textos gerados por IA tendem a apresentar menor variação de palavras e sintática, outra métrica utilizada é a “burstiness”, que mede a inconsistência ou variação no textos de produções humanas

(IPPOLITO et al., 2020, s.p.). Com base no mencionado, quais são as novas ferramentas que estão sendo mais utilizadas na detecção de texto gerados por IAs e como elas particularmente funcionam.

2.5.4.1 Ferramenta de Detecção de IA em textos (GPTZero)

Desenvolvido por Edward Tian em 2023, foi um dos pioneiros na detecção de IA generativa em textos. Seu compromisso é garantir a distinção dos textos de autoria humana dos textos gerados por LLM, para cumprir esse objetivo o GPTzero oferece um detector capaz de fornecer previsões explicáveis que permitem os usuários interpretar os dados e chegarem a uma conclusão própria com responsabilidade (Tian, 2023).

No GPTZero é utilizado o Deep Learning (aprendizado profundo) de ponta a ponta, treinada exclusivamente para detectar conjuntos de dados frequentemente usados por diversos LLMs; Com base nesses conjuntos de frase ele determina a probabilidade e a confiança de que um texto foi criado por IA (Tian, 2023).

O GPTZero oferece aplicações em diversas áreas, como verificador de plágio acadêmico para verificar em trabalhos ou avaliações se foi utilizado alguma IA para realizar as tarefas sem a devida autoria ou para descobrir de quem é o real autor de um texto prezando a proteção de seus direitos autorais (Tian, 2023).

2.5.4.2 Ferramenta de Detecção de IA em textos (ZeroGPT)

A ferramenta de detecção de ia em textos ZeroGPT utiliza os próprios padrões de detecção das outras IA's como ChatGPT da OpenIA, Gemini da google entre outras IAs para moldar a estrutura de frases e repetições de palavras mais utilizadas pelas inteligências artificiais para detectar essas padrões nos textos a serem detectados (ZeroGPT, 2025).

O grande diferencial do ZeroGPT é sua interface simples e autoexplicativa que facilita o uso do detector para usuários leigos as suas vantagens são suas atualizações frequentes para acompanhar os novos modelos de IA se mantendo atualizados com as constantes mudanças do mercado de inteligência artificial (ZeroGPT, 2025).

2.5.4.3 Ferramenta de Detecção de IA em textos (Originality.AI)

O Originality.AI é uma ferramenta de detecção de IA desenvolvida para identificar com grande precisão os textos gerados por modelos de IA como GPT-3 e GPT-4. Seu diferencial é a combinação do machine learning e uma lógica avançada que permite uma análise mais detalhada e complexa dos textos para poder determinar sua origem (Originality.AI, 2025).

Alguns de seus focos são a prevenção de plágios, fraudes acadêmicas e garantir a autenticidade em artigos jornalísticos e profissionais (Originality.AI, 2025).

2.5.5 Discussão crítica

A partir disso podemos notar que os detectores de IA são desenvolvidos com alguns objetivos em comuns como a detecção eficaz de texto gerados por LLM's, detecção de plágio e autoria. Embora sua utilidade podemos notar que são ferramentas limitadas. Entre os diversos problemas estão as taxas de falsos positivos/negativos que podem prejudicar alunos/usuários injustamente os autores causando dilemas éticos.

Todas as ferramentas de verificações apresentam perspectivas promocionais que informam que a mesma é a mais precisa, confiável e capaz de identificar com precisão os textos gerados por inteligências artificiais, em alguns casos possuem uma análise comparativa com outras ferramentas de verificação de IA visando alegar uma precisão maior e melhores avaliações dos usuários.

Visando verificar a autenticidade dessas alegações e a análise de quais são as

ferramentas de verificação de IA com as melhores precisões, foi feita a análise comparativa dessas ferramentas selecionadas, conforme na partição metodologia.

3. METODOLOGIA

A presente seção detalha o delineamento técnico empregado para avaliar o desempenho das ferramentas de detecção de Inteligência Artificial (IA) no contexto de textos jornalísticos. A metodologia é de natureza experimental e comparativa, com abordagem predominantemente quantitativa, visando a mensuração objetiva e a análise estatística do desempenho dos detectores.

3.1 Tipo de Pesquisa e Abordagem

Este estudo configura-se como uma pesquisa de natureza predominantemente experimental e comparativa, com abordagem quantitativa na análise dos dados. O caráter exploratório justifica-se pela ausência de um padrão consolidado na avaliação de ferramentas de detecção de IA, enquanto o aspecto experimental decorre da criação controlada de textos em um ambiente delimitado, permitindo observar empiricamente como essas ferramentas se comportam diante de diferentes origens textuais : humana, artificial e híbrida. A abordagem quantitativa será empregada na mensuração objetiva do desempenho dos detectores, por meio de dados numéricos e índices estatísticos, ao passo que a comparativa possibilitará estabelecer as capacidades relativas e as limitações de cada ferramenta. Por fim, o estudo possui caráter aplicado, visto que seus resultados poderão contribuir diretamente para a compreensão e o aprimoramento das práticas de detecção de IA em contextos acadêmicos, jornalísticos e institucionais, incentivando o uso ético e transparente dessas tecnologias emergentes, conforme apontado na introdução deste trabalho.

3.2 Objeto de Estudo e Ferramentas Empregadas

O objeto central deste estudo reside na avaliação crítica do desempenho de ferramentas de detecção de Inteligência Artificial (IA) frente a textos gerados com diferentes níveis de intervenção humana. Para tal, foi selecionada uma ferramenta de IA generativa e três detectores de IA.

A inteligência artificial utilizada para a geração de textos é o ChatGPT, na sua versão GPT-5. Sua escolha justifica-se pela sua ampla notoriedade e utilização global, conforme evidenciado na seção 2.3.1 deste trabalho, que aponta para uma base de 800 milhões de usuários ativos semanais em agosto de 2025. O que torna o ChatGPT uma ferramenta representativa das tecnologias de geração de linguagem natural mais avançadas e amplamente empregadas atualmente.

Para a detecção dos textos, foram selecionadas três ferramentas comumente citadas no cenário atual: GPTZero, ZeroGPT e Originality.ai. Esta seleção buscou garantir uma diversidade no escopo de testes, incluindo dois detectores de acesso gratuito (GPTZero e ZeroGPT) e um de caráter pago (Originality.ai), o que permite uma análise mais abrangente das soluções disponíveis no mercado. Um critério fundamental para a escolha destas ferramentas, e para assegurar a relevância dos testes, foi a verificação de seu suporte à detecção de conteúdos gerados pela versão GPT-5 do ChatGPT e ao processamento de textos na língua portuguesa. Esse alinhamento garante um parâmetro justo e pertinente para as análises comparativas.

3.3 Delineamento Experimental e Geração de Dados

O delineamento experimental da pesquisa será conduzido mediante a submissão de um conjunto de doze textos às três ferramentas de detecção de IA previamente selecionadas. É crucial ressaltar que todos esses textos, independentemente de sua autoria, pertencem ao gênero jornalístico, especificamente do tipo notícia, mantendo a uniformidade temática e estrutural para a avaliação. Esses doze textos estão equitativamente distribuídos em três categorias de autoria: 4 textos gerados exclusivamente por humanos (100% humanos), 4 textos gerados

integralmente por inteligência artificial (100% IA) e 4 textos de natureza híbrida. Cada um dos detectores será, consequentemente, testado com cada um desses doze textos, resultando em um total de 36 análises individuais. Os testes serão sistematicamente executados em um ambiente de navegação limpo, destituído de histórico de pesquisas prévias, e os resultados obtidos serão considerados válidos para a data específica de realização das análises, fixada em outubro de 2025.

Para uma compreensão sintetizada da configuração experimental, a Tabela 6 apresenta um panorama das principais variáveis envolvidas no delineamento.

Tabela 6 - Configuração do Delineamento Experimental

Variável Experimental	Detalhe da Configuração
Número Total de Textos Analisados	12
Categorias de Autoria dos Textos	Humana (4 textos), 100% IA (4 textos) e Híbrida (4 textos)
Gênero Textual dos Textos	Jornalístico (Notícia)
Ferramentas de Detecção Avaliadas	GPTZero, ZeroGPT, Originality.ai (3 ferramentas)
Número Total de Análises Individuais	36 (12 textos x 3 detectores)
Período de Validade dos Resultados	Outubro de 2025

Fonte : Elaborada pelos autores

3.3.1 Critérios de Seleção e Geração dos Textos

Para garantir a validade e a representatividade da amostra textual, a geração e seleção dos textos seguirão os seguintes critérios:

3.3.1.1 Textos Gerados por Humanos (100% Humanos)

Para esta categoria, foram cuidadosamente selecionados quatro textos do tipo notícia, publicados em portais de notícias brasileiros antes do dia 22 de novembro de 2022. Esta delimitação temporal é estratégica e crucial, pois antecede o lançamento público do ChatGPT e de outras Large Language Models (LLMs), garantindo que os conteúdos sejam integralmente originais e livres de qualquer possível intervenção ou influência de tecnologias de IA generativa. Todos os textos selecionados apresentam um comprimento mínimo de 500 palavras e um mínimo de 3.000 caracteres, conforme contagem verificada pela ferramenta online invertexto.com, garantindo uma extensão adequada e consistente para a análise pelos detectores. Visando uma amostragem que reflete a riqueza de estilos e abordagens dentro do gênero notícia, foram escolhidos artigos que representam as tipologias descritas na Tabela 7.

Tabela 7 - Tipologias de Notícia

Tipo de Notícias	Descrição
Análise Crítica	Textos que, além de apresentar informações, oferecem uma interpretação aprofundada ou um posicionamento sobre a situação abordada.
Notícia Insólita	Artigos que relatam eventos ou fatos incomuns, estranhos ou extraordinários, mas cuja veracidade é comprovável.
Notícia Factual	Reportagens que se restringem à exposição objetiva e imparcial dos fatos, sem a inserção de opiniões ou da defesa de ideologias.
Notícia de Dados	Conteúdos que se baseiam na apresentação e interpretação de dados, números e levantamentos estatísticos, frequentemente originários de fontes oficiais e verificadas.

Fonte : Elaborado pelos autores.

3.3.1.2 Textos Gerados integralmente por Inteligência Artificial

Para esta categoria, foram gerados quatro novos textos correspondendo a cada um dos temas dos textos humanos originais, porém produzidos integralmente pela IA ChatGPT (versão GPT-5). A inclusão desta categoria de textos teve como objetivo estabelecer um ponto de partida irrefutável para a avaliação das ferramentas de detecção. Esses textos representam cenários de autoria puramente sintética, onde a intervenção humana se limitou à formulação do comando inicial. Isso permite aferir a capacidade dos detectores em identificar a presença inequívoca de conteúdo gerado por IA, sem qualquer edição ou aprimoramento humano posterior, servindo como um exemplo puro de texto artificial para comparação.

A geração de cada texto foi realizada no ChatGPT (versão GPT-5) empregando um prompt padrão cuidadosamente elaborado. A estrutura do prompt foi concebida para simular uma solicitação profissional e detalhada, segurando que IA produzisse um conteúdo pertinente e dentro do escopo do estudo. O modelo completo do prompt padrão utilizado é apresentado no Apêndice A. Os elementos que compõem este prompt padrão são detalhados na Tabela 8:

Tabela 8 - Elementos que compõem do prompt

Elemento do Prompt	Descrição	Objetivo
Persona	A IA foi instruída a atuar como um "redator profissional de jornal especializado em política e economia brasileira".	Direcionar o estilo e o vocabulário da IA para um perfil autêntico de produção jornalística, simulando uma solicitação profissional real para obter resultados mais verossímeis.
Linguagem e Tom	Exigência de linguagem jornalística imparcial, clara, objetiva e sem adjetivos opinativos. A menção a fontes oficiais ou referências conhecidas também foi incluída.	Fundamental para alinhar os textos gerados ao escopo do estudo (textos jornalísticos) e para garantir a alta qualidade e pertinência do conteúdo produzido para a análise.

Extensão	Indicação de aproximadamente 500 palavras, com a condição explícita de que o texto final deveria estar "pronto para publicação".	Reforçar a expectativa de um produto final coeso, completo e com qualidade editorial, além de padronizar a quantidade de texto para análise.
----------	--	--

Fonte: Elaborado pelos autores.

Além das diretrizes gerais de persona, linguagem e extensão, o prompt especificou detalhadamente a composição estrutural da notícia, garantindo que a IA produzisse textos com um formato jornalístico padronizado. A Tabela 9 detalha os componentes exigidos para cada notícia gerada:

Tabela 9 - Elementos que compõem do prompt

Componente da Notícia	Descrição	Requisitos do Prompt
Título Principal	Frase ou expressão inicial que apresenta o tema principal.	Deveria ser chamativo e informativo, capturando a essência do conteúdo da notícia.
Subtítulo (Linha Fina)	Texto breve que complementa o título.	Com uma função de contextualização breve e complementar ao título, expandindo a informação inicial e preparando o leitor para o tema.
Lead	Parágrafo inicial que resume os fatos essenciais.	Necessidade de resposta concisa e direta aos elementos jornalísticos essenciais: O Quê, Quem, Quando, Onde, Como e Por Quê, em conformidade com as boas práticas do jornalismo.
Corpo da Notícias	Desenvolvimento detalhado dos fatos e análises.	Exigido em formato detalhado, incluindo contexto histórico, repercussão dos fatos, potenciais citações de fontes (mesmo que genéricas) e análise de impactos (econômicos, sociais, políticos, etc.), para fornecer profundidade e completude.

Conclusão	Fechamento da narrativa com síntese ou desdobramentos.	Orientada para apresentar uma análise final, projeções de desdobramentos futuros ou uma síntese do impacto do evento, fechando a narrativa de forma coesa.
-----------	--	--

Fonte: Elaborado pelos autores.

Para garantir que uma quantidade e tipo similar de informação fossem fornecidos à IA para cada geração, evitando assim influenciar os resultados, o prompt incluiu uma estrutura padronizada para a entrada de dados base da notícia:

“{[Assunto principal], assegura [principais autoridade do assunto se tiver] , [houve auxílio de : principais atores/instituições], ocorrido em [data ou ano ocorrido], analisando [seus impactos/desdobramentos principais]}”

Essa estrutura foi preenchida com informações específicas baseadas nas notícias humanas originais correspondentes (**Anexos A, B, C e D**). O objetivo foi assegurar que a mesma quantidade de informação contextual fosse fornecida para a IA na geração dos textos, evitando vieses e influências indevidas na amostragem que pudessem comprometer a comparação entre os textos.

A Tabela 10 ilustra como essa parte do prompt foi detalhada para cada tema, resultando na geração dos textos que compõem os **Apêndices C, D, E e F**:

Tabela 10 - Formato de preenchimento de prompt

Anexo usado como referência	Formato de preenchimento de prompt
Anexo A	O debate global acerca da privacidade digital e a iminente adoção da LGPD no Brasil, analisado no contexto dos escândalos de Edward Snowden e Cambridge Analytica, envolvendo a competição geopolítica de dados entre China e EUA, e a busca por regulamentação no Ocidente, referente aos escândalos de 2013 e 2016 e à entrada em vigor da LGPD em agosto de 2020, e que explore seus impactos na proteção da democracia, na soberania digital e na nova economia de consultoria em dados.

Anexo B	A surpreendente capacidade de cabras de discriminar expressões faciais humanas, divulgada em estudo na revista Royal Society Open Science por pesquisadores internacionais, envolvendo as cientistas brasileiras Natália de Souza Albuquerque e Carine Savalli e o Santuário de Buttercups, referente a um estudo publicado em 2018 que explore seus questionamentos sobre a cognição animal, a origem evolutiva da percepção de emoções e a complexidade das relações humano-animais.
Anexo C	As novas exigências de dados de redes sociais para solicitantes de visto americano, anunciado pelo Departamento de Estado americano, envolvendo a determinação do então presidente Donald Trump, a maioria dos viajantes (turismo, estudo, trabalho), a União Americana pelas Liberdades Civis (ACLU) e usuários das redes sociais, referente à proposta original em março de 2018 e ao início da implementação em 2019, com solicitação de dados dos últimos cinco anos, e que explore seus impactos no aumento do controle de segurança, as preocupações com a privacidade e a autocensura online, as críticas à eficácia e justiça da medida, e as possíveis consequências para quem fornecer informações falsas.
Anexo D	A projeção do crescimento e envelhecimento da população brasileira, analisado na revisão 2018 da Projeção de População do IBGE, envolvendo a migração venezuelana em Roraima e as expectativas de queda da fecundidade por estados, com dados projetados anualmente até 2060, a partir da revisão de 2018, e que explore as implicações para o mercado de trabalho, a previdência social e a estrutura socioeconômica do país.

Fonte : Elaborado pelos autores.

3.3.1.3 Textos Híbridos

Esta categoria compreende quatro textos jornalísticos que resultaram da colaboração entre autoria humana e aprimoramento por Inteligência Artificial. O objetivo principal da criação desses textos híbridos foi simular cenários mais realistas de produção de conteúdo, onde a IA atua como uma ferramenta de apoio editorial, refinando e otimizando textos originalmente humanos, mas sem reescrevê-los ou alterar seu conteúdo factual essencial. Esta abordagem permite investigar a capacidade das ferramentas de detecção em identificar nuances de aprimoramento por IA, mantendo a base autoral humana.

Como citado anteriormente em produção textual, foram considerados como textos híbridos todos aqueles que sofreram influência de um sistema de Inteligência Artificial durante seu processo de produção.

Optou-se por essa definição ampla em razão da impossibilidade de estabelecer os níveis exatos de influência de IA e de contribuição humana em cada texto. A ausência de métricas precisas que permitam quantificar o grau de intervenção artificial torna inviável uma classificação mais quantitativa, conduzindo à adoção de uma categorização binária: textos que receberam qualquer tipo de assistência de IA são classificados como híbridos.

O processo de geração dos textos híbridos envolveu a utilização dos quatro textos humanos originais (**Anexos A, B, C e D**) como base. Cada um desses textos foi submetido à IA ChatGPT (versão GPT-5) com um prompt específico, desenvolvido para instruir a IA a realizar um trabalho de edição e aprimoramento. O resultado desse processo de aprimoramento por IA são os textos híbridos, que foram disponibilizados integralmente nos **Apêndices G, H, I e J**.

O modelo completo do prompt utilizado para a geração dos textos híbridos é apresentado no **Apêndice B**. As diretrizes gerais e a configuração inicial do prompt para os textos híbridos são detalhadas na Tabela 11:

Tabela 11 - Formato de preenchimento de prompt

Elemento do Prompt	Descrição	Racional Estratégico
Persona	A IA foi instruída a atuar como um "editor de notícias experiente e refinado, especializado em política e economia brasileira".	Essa persona foi escolhida para simular um papel profissional de revisão e aprimoramento, alinhado ao gênero jornalístico do estudo, buscando uma IA que compreendesse as sutilezas da edição de notícias.

Objetivo Geral	"Aprimorá-lo para publicação, garantindo que ele seja acessível, envolvente e de fácil compreensão para um público amplo, ao mesmo tempo em que mantém a precisão dos fatos e a integridade do conteúdo original."	Este objetivo estratégico diferencia os textos híbridos dos textos integralmente gerados, pois foca na otimização de um texto preexistente, não em sua criação do zero. É crucial para testar a detecção de IA em cenários de edição sutil e controlada.
Base do Aprimoramento	Indicação explícita para inserir o texto original humano que deveria ser editado.	Define a base para a atuação da IA como editora, vinculando diretamente o processo de aprimoramento ao material de origem humana e garantindo que a IA trabalhe sobre conteúdo preexistente.

Fonte : Elaborado pelos autores.

Complementarmente às diretrizes gerais, o prompt forneceu instruções específicas à IA para apenas o aprimoramento do texto, garantindo que as modificações fossem realizadas dentro de parâmetros controlados, focando na fluidez, clareza e manutenção da integridade do conteúdo original. Essas instruções detalhadas são apresentadas na Tabela 12:

Tabela 12 - Instrução Específica de Aprimoramento

Instrução	Descrição	Objetivo
1 - Fluidez e Clareza Narrativa	"Reorganize a ordem dos fatos, se necessário, para garantir a máxima fluidez e uma sequência lógica que torne a leitura fácil e comprehensível. O texto deve progredir de forma natural, guiando o leitor através das informações sem obstáculos."	Visava a melhoria da estrutura e coesão textual, características frequentemente associadas a uma boa edição jornalística, sem alterar o conteúdo factual principal.

2- Vocabulário e Acessibilidade	"Avalie o vocabulário utilizado e substitua termos menos comuns, jargões ou complexos por sinônimos mais acessíveis e diretos, sem comprometer o sentido ou o tom profissional da notícia. O objetivo é que o texto seja facilmente digerível por um leitor sem conhecimento prévio aprofundado no tema."	Focou na simplificação da linguagem para aumentar o alcance e a compreensão do público, um trabalho típico de um editor profissional.
3 - Manutenção de Citações e Fontes	"Todas as citações de fontes, declarações e referências presentes no texto original devem ser mantidas exatamente como estão, sem quaisquer alterações em seu conteúdo ou atribuição."	Esta instrução é de suma importância estratégica. Garante a preservação da autoria humana e da integridade factual, reforçando que a IA atua como editora e não como autora primária.
4 - Estilo Jornalístico Profissional	"Assegure que o texto final preserve um tom jornalístico imparcial, claro e objetivo. O aprimoramento deve resultar em um texto que seja polido, profissional e natural, como se tivesse passado por uma revisão editorial de um grande portal de notícias (exemplo: G1, Estadão, Folha de S.Paulo)."	Garantir que as melhorias aplicadas pela IA resultam em um texto com a qualidade editorial esperada de uma publicação profissional, mantendo a autenticidade do estilo jornalístico, conferindo um acabamento polido e natural, condizente a uma revisão
5 -Polimento Geral	"Revise a gramática, sintaxe e pontuação para eliminar quaisquer imperfeições, garantindo que o texto esteja impecável para publicação."	Cobre os aspectos finais de revisão, comuns a qualquer processo editorial, garantindo a correção formal e a qualidade final do texto.

Fonte: Elaborado pelos autores.

3.3.2 Resumo da Amostra Textual

Para facilitar a compreensão e referência dos textos utilizados ao longo do experimento, a Tabela 13 apresenta um resumo dos doze documentos que comporão a amostra, indicando sua categoria autoral, o tipo de notícia (conforme definido na Tabela 2) e uma breve descrição do conteúdo. Os textos completos, juntamente com os prompts utilizados para sua geração, serão disponibilizados nos apêndices deste trabalho.

Tabela 13 - Resumo dos Textos Componentes da Amostra para Análise.

ID do Texto	Categoria	Tipo de Notícias	Descrição Breve	Referência no Apêndice
H1	Humano	Crítica	O artigo discute os impactos da coleta de dados e vigilância global.	Anexo A
H2	Humano	Insólita	Estudo mostra que cabras conseguem identificar emoções em rostos humanos.	Anexo B
H3	Humano	Factual	Novas exigências de visto americano obrigam os viajantes a fornecer dados de redes sociais, gerando preocupações sobre privacidade e autocensura.	Anexo C
H4	Humano	Dados	Estudo prevê que a população brasileira crescerá até 2047 e depois começará a diminuir, com envelhecimento acentuado.	Anexo D
IA1	100% IA	Crítica	Gerado por IA com base no tema de H1.	Apêndice C
IA2	100% IA	Insólita	Gerado por IA com base no tema de H2.	Apêndice D

IA3	100% IA	Factual	Gerado por IA com base no tema de H3.	Apêndice E
IA4	100% IA	Dados	Gerado por IA com base no tema de H4.	Apêndice F
HYB1	Híbrido	Crítica	Texto H1 aprimorado/reescrito por IA.	Apêndice G
HYB2	Híbrido	Insólita	Texto H2 aprimorado/reescrito por IA.	Apêndice H
HYB3	Híbrido	Factual	Texto H3 aprimorado/reescrito por IA.	Apêndice I
HYB4	Híbrido	Dados	Texto H4 aprimorado/reescrito por IA.	Apêndice J

Fonte : Elaborado pelos autores.

3.3.3 Ambiente de Teste Controlado e Delimitação Temporal

Para garantir a validade e a consistência dos resultados, a geração de todos os textos via ChatGPT (GPT-5) foi realizada em um ambiente de teste controlado. Para cada texto gerado, foi utilizada uma conta "limpa" e sem histórico de interações prévias, minimizando a influência de vieses contextuais. Após a geração de cada texto, o histórico de conversas no ChatGPT foi imediatamente apagado para evitar qualquer tipo de "memória" ou adaptação do modelo que pudesse comprometer a neutralidade das gerações subsequentes.

Os textos foram gerados no dia 16/10/2025, e as análises pelos detectores foram conduzidas entre 18/10/2025 e 26/10/2025. É crucial ressaltar que os resultados desta pesquisa refletem o estado da arte das ferramentas de IA generativa e de detecção de IA até o período da coleta de dados. Dada a velocidade vertiginosa de evolução dessas tecnologias, os resultados obtidos são intrinsecamente restritos ao momento da execução do experimento, e futuras atualizações dos modelos de IA ou dos detectores podem alterar significativamente os padrões de identificação.

3.4 Coleta e Análise de Dados

Após a meticulosa geração e categorização dos doze textos (quatro de cada tipo), cada um deles será submetido individualmente a cada uma das três ferramentas de detecção de IA (GPTZero, ZeroGPT e Originality.ai). Para cada submissão, serão registrados os resultados fornecidos pelos detectores, que tipicamente indicam uma probabilidade ou classificação de autoria (humana, IA ou híbrida) e, em alguns casos, destacam trechos específicos.

A análise dos dados será realizada de maneira quantitativa, avaliando as seguintes métricas de desempenho para cada detector:

1. **Precisão:** Será calculada a proporção, em porcentagem, do total de classificações corretas realizadas por cada detector em relação ao seu tipo de autoria real (humano, IA ou híbrido).
2. **Taxa de Falsos Positivos:** Será mensurado o percentual de textos comprovadamente humanos que foram erroneamente classificados como gerados por IA pelos detectores. Esta métrica é crucial para avaliar a confiabilidade da ferramenta ao não acusar indevidamente a autoria humana.
3. **Taxa de Falsos Negativos:** Será determinado o percentual de textos comprovadamente gerados por IA (ou com intervenção significativa de IA) que foram incorretamente classificados como humanos pelos detectores. Esta métrica é vital para verificar a capacidade do detector em identificar a autoria artificial.

Tabela 14 - Tabela de termos

Termo	Significado
VP (Verdadeiro Positivo)	Texto de IA corretamente identificado como IA

VN (Verdadeiro Negativo)	Texto humano corretamente identificado como humano
FP (Falso Positivo)	Texto humano incorretamente identificado como IA
FN (Falso Negativo)	Texto de IA incorretamente identificado como humano

Fonte : Elaborado pelos autores

Conforme demonstrado na tabela 14, os cálculos serão elaborados utilizando equações lineares para a análises das taxas de precisão, falso positivo e falso negativo.

Figura 8 - Fórmulas de Equação linear utilizadas

$$\text{Precisão} = \frac{VP+VN}{FP+FN+VP+VN}$$

$$\text{Taxa de falso positivo} = \frac{FP}{FP + VN}$$

$$\text{Taxa de falso negativo} = \frac{FN}{FN + VP}$$

Fonte : Elaborado pelos autores.

Essas fórmulas foram utilizadas para definir o quanto o detector acerta ao identificar trechos humanos e textos gerados por IA.

Para a tabulação, organização e visualização dos resultados, serão utilizadas

ferramentas como o Microsoft Excel . Essas ferramentas auxiliarão na elaboração de gráficos comparativos que ilustrarão a precisão, a quantidade de acertos e erros de cada detector, proporcionando uma simples e eficaz identificação dos verificadores com maior eficácia no contexto dos textos analisados.

3.5 Limitações da Pesquisa

É fundamental reconhecer as limitações inerentes a este estudo para contextualizar adequadamente seus resultados. Primeiramente, o tamanho da amostra de teste, composta por doze documentos, embora suficiente para uma análise comparativa inicial, pode não ser representativo de toda a diversidade e complexidade das produções textuais encontradas em cenários reais. Em segundo lugar, a pesquisa abrange apenas um modelo de Inteligência Artificial (ChatGPT GPT-5) para a elaboração dos textos e um conjunto específico de três detectores de IA. Consequentemente, os resultados obtidos não podem ser generalizados para todas as ferramentas de IA generativa ou para todos os detectores disponíveis no mercado.

Outro ponto crucial é a constante e acelerada evolução tanto das tecnologias de IA generativa quanto dos próprios detectores. Essa dinâmica contínua pode levar a variações nos resultados em futuras replicações do mesmo experimento; assim, os resultados obtidos são um retrato do desempenho dessas ferramentas no período de 18/10/2025 a 26/10/2025, e podem não se manter válidos em um futuro próximo devido às constantes atualizações e aprimoramentos tecnológicos. Diante desses fatos, reitera-se que este estudo se trata de uma análise exploratória e experimental que, com base no cenário aplicado e no período testado, oferece insights valiosos para a compreensão da eficácia dos detectores de IA, pavimentando o caminho para trabalhos futuros mais abrangentes.

4. ANÁLISE DOS RESULTADOS

A precisão constitui o principal critério de avaliação neste estudo, sendo calculada com base na quantidade de classificações corretas realizadas por cada detector, de acordo com cada texto (humano, IA ou híbrido).

Para cada ferramenta, foram analisadas as taxas de acertos, falsos positivos (textos humanos classificados como IA) e falsos negativos (textos de IA classificados como humanos). A Tabela 15 em diante apresentam as análises dos resultados obtidos.

Tabela 15 - Análise Texto 100% Humano 1 - H1/Crítica

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	33,41	66,59	Não fornece esse parâmetro
GPTZero	0,00	96,00	4,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte : Elaborado pelos autores.

Os três verificadores apresentaram divergências significativas na classificação deste texto humano. O ZeroGPT identificou 66,59% como conteúdo humano, demonstrando desempenho intermediário. O GPTZero obteve o melhor resultado entre as ferramentas, classificando 96% como humano e apenas 4% como híbrido, aproximando-se da origem real do texto. Em contrapartida, o Originality.AI apresentou um falso positivo severo, atribuindo 100% de IA ao texto genuinamente humano, evidenciando limitações na detecção desta ferramenta para este tipo específico de conteúdo.

Tabela 16 - Análise Texto 100% Humano 2 - H2/Insólita

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	46,51	53,49	Não fornece esse parâmetro

GPTZero	0,00	100,00	0,00
Originality	0,00	100,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores.

Os três verificadores demonstraram desempenho mais consistente neste texto humano em comparação com o anterior. O ZeroGPT classificou 53,49% como humano e 46,51% como IA. O GPTZero e o Originality.AI obtiveram desempenho ideal, ambos identificando 100% do conteúdo como humano e 0% como IA, demonstrando concordância total na detecção.

Tabela 17 - Análise Texto 100% Humano 3 - H3/Factual

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	67,24	32,76	Não fornece esse parâmetro
GPTZero	0,00	100,00	0,00
Originality	0,00	100,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores.

Observa-se divergência significativa entre os verificadores neste texto de caráter factual. O ZeroGPT apresentou falso positivo considerável, classificando 67,24% do conteúdo como IA e apenas 32,76% como humano, invertendo a natureza real do texto. Em contrapartida, o GPTZero e o Originality.AI mantiveram o desempenho ideal, ambos identificando 100% como humano e 0% como IA.

Tabela 18 - Análise Texto 100% Humano 4 - H4/Dados

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	92,85	7,15	Não fornece esse parâmetro

GPTZero	4,00	89,00	7,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores.

Os três verificadores apresentaram divergências acentuadas na classificação deste texto humano baseado em dados. O ZeroGPT demonstrou falso positivo, atribuindo 92,85% a IA e apenas 7,15% ao humano. O GPTZero obteve melhor desempenho relativo, classificando 89% como humano, 4% como IA e 7% como híbrido, aproximando-se da origem real do texto. O Originality.AI apresentou falso positivo, identificando 100% como IA.

Tabela 19 - Análise Texto 100% IA 1 - IA1/Crítica

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	100,00	0,00	Não fornece esse parâmetro
GPTZero	82,00	6,00	12,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores.

Os três verificadores demonstraram alta eficácia na identificação deste texto gerado integralmente por IA. O ZeroGPT e o Originality.AI obtiveram desempenho perfeito, ambos classificando 100% do conteúdo como IA e 0% como humano, detectando corretamente a origem artificial do texto. O GPTZero também apresentou resultado próximo da realidade, atribuindo 82% a IA, 6% a humano e 12% como híbrido, embora com maior margem de incerteza em comparação às outras ferramentas.

Tabela 20 - Análise Texto 100% IA 2 - IA2/Insólita

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
------------	--------	------------	-------------

ZeroGPT	78,55	21,45	Não fornece esse parâmetro
GPTZero	79,00	18,00	3,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores identificaram predominantemente a origem artificial deste texto, embora com variações nos percentuais. O ZeroGPT classificou 78,55% como IA e 21,45% como humano. O GPTZero apresentou resultado similar, atribuindo 79% a IA, 18% a humano e 3% como híbrido, mantendo convergência com o ZeroGPT. O Originality.AI obteve desempenho perfeito, identificando 100% como IA e 0% como humano.

Tabela 21 - Análise Texto 100% IA 3 - IA3/Factual

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	89,37	10,63	Não fornece esse parâmetro
GPTZero	54,00	0,00	46,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores identificaram corretamente a origem artificial deste texto factual, porém com níveis variados de confiança. O ZeroGPT classificou 89,37% como IA e 10,63% como humano. O GPTZero apresentou resultado interessante, atribuindo 54% a IA, 0% a humano e 46% como híbrido, reconhecendo características mistas no texto. O Originality.AI manteve desempenho perfeito, identificando 100% como IA.

Tabela 22 - Análise Texto 100% IA 4 - IA4/Dados

Ferramenta	IA (%)	Humano (%)	Híbrido (%)

ZeroGPT	98,66	1,34	Não fornece esse parâmetro
GPTZero	83,00	9,00	8,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores demonstraram alta eficácia na identificação deste texto gerado por IA com apresentação de dados. O ZeroGPT obteve desempenho quase perfeito, classificando 98,66% como IA e apenas 1,34% como humano, evidenciando forte detecção da origem artificial. O GPTZero também apresentou resultado satisfatório, atribuindo 83% a IA, 9% a humano e 8% como híbrido, mantendo a identificação predominante de conteúdo artificial. O Originality.AI alcançou desempenho ideal, identificando 100% como IA e 0% como humano.

Tabela 23 - Análise Texto Híbrido 1 - HYB1/Crítica

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	43,47	56,53	Não fornece esse parâmetro
GPTZero	83,00	12,00	5,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores apresentaram divergências acentuadas na classificação deste texto híbrido. O ZeroGPT identificou 56,53% como humano e 43,47% como IA, demonstrando equilíbrio nos percentuais que reflete parcialmente a natureza mista do conteúdo, embora com leve tendência à classificação humana. O GPTZero atribuiu 83% a IA, 12% a humano e apenas 5% como híbrido, inclinando-se fortemente para a detecção de conteúdo artificial e subutilizando sua categoria híbrida. O Originality.AI

classificou 100% como IA, desconsiderando completamente a contribuição humana no texto.

Tabela 24 - Análise Texto Híbrido 2 - HYB2/Insólita

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	54,17	45,83	Não fornece esse parâmetro
GPTZero	7,00	92,00	1,00
Originality	99,00	1,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores apresentaram resultados completamente divergentes na classificação deste texto híbrido de estilo insólito. O ZeroGPT identificou 54,17% como IA e 45,83% como humano, demonstrando equilíbrio nos percentuais. O GPTZero apresentou resultado oposto, atribuindo 92% a humano, 7% a IA e apenas 1% como híbrido, classificando o texto predominantemente como humano e ignorando sua categoria híbrida. O Originality.AI classificou 99% como IA e 1% como humano, inclinando-se quase totalmente para detecção artificial.

Tabela 25 - Análise Texto Híbrido 3 - HYB3/Factual

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	62,21	37,79	Não fornece esse parâmetro
GPTZero	39,00	57,00	4,00
Originality	99,00	1,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores apresentaram divergências significativas na classificação deste texto híbrido factual. O ZeroGPT identificou 62,21% como IA e 37,79% como humano, demonstrando distribuição que reconhece a presença de ambos os elementos, com leve predominância de detecção artificial. O GPTZero atribuiu 57% a humano, 39% a IA e 4% como híbrido, inclinando-se discretamente para classificação humana e apresentando percentuais relativamente equilibrados. O Originality.AI classificou 99% como IA e apenas 1% como humano, desconsiderando quase completamente a contribuição humana no texto.

Tabela 26 - Análise Texto Híbrido 4 - HYB4/Dados

Ferramenta	IA (%)	Humano (%)	Híbrido (%)
ZeroGPT	90,26	9,74	Não fornece esse parâmetro
GPTZero	14,00	46,00	19,00
Originality	100,00	0,00	Não fornece esse parâmetro

Fonte: Elaborado pelos autores

Os três verificadores apresentaram divergências acentuadas na classificação deste texto híbrido com apresentação de dados. O ZeroGPT identificou 90,26% como IA e apenas 9,74% como humano, classificando o texto predominantemente como artificial e desconsiderando amplamente a contribuição humana. O GPTZero atribuiu 46% a humano, 14% a IA e 19% como híbrido, sendo o único texto onde a ferramenta utilizou significativamente sua categoria híbrida, embora ainda com predominância da classificação humana. O Originality.AI classificou 100% como IA, mantendo seu padrão de detecção categórica observado nos demais textos híbridos. Apenas o GPTZero apresentou distribuição que reconhece características mistas através do uso relevante da categoria híbrida (19%), enquanto ZeroGPT e Originality.AI inclinaram-se fortemente para classificação artificial.

Tabela 27 - Média porcentagem dos textos 100 % humanos - H1,H2,H3 E H4

Ferramenta	Humano (%)	IA (%)
ZeroGPT	39,747	60,25
GPTZero	96,25	3,75
Originality	50,00	50,00
Média Geral das ferramentas:	62	38

Fonte: Elaborado pelos autores

Esta tabela consolida o desempenho médio dos três verificadores na análise dos quatro textos genuinamente humanos. O ZeroGPT identificou 39,747% como humano e 60,25% como IA, apresentando maior tendência a classificar conteúdo humano como artificial. O GPTZero classificou 96,25% como humano e 3,75% como IA, aproximando-se da origem real dos textos. O Originality.AI apresentou 50% para ambas as categorias, indicando equilíbrio entre as classificações. A média geral das três ferramentas resultou em 62% humano e 38% IA.

Tabela 28 - Média porcentagem dos textos 100 % IA - IA1,IA2,IA3 E IA4

Ferramenta	Humano (%)	IA (%)
ZeroGPT	8,355	91,645
GPTZero	8,25	91,75
Originality	0	100
Média Geral das ferramentas:	5,54	94,46

Fonte: Elaborado pelos autores

Esta tabela consolida o desempenho médio dos três verificadores na análise dos quatro textos gerados integralmente por IA. O ZeroGPT classificou 91,645% como IA e 8,355% como humano, demonstrando alta capacidade de detecção de conteúdo artificial. O GPTZero apresentou resultado similar, atribuindo 91,75% a IA e 8,25% humano, mantendo convergência com o ZeroGPT. O Originality.AI obteve desempenho perfeito, identificando 100% como IA e 0% como humano em todos os textos analisados. A média geral das três ferramentas resultou em 94,46% de IA e 5,54% humano.

Tabela 29 - Média porcentagem dos híbridos - HYB1, HYB2, HYB3 E HYB4

Ferramenta	Humano (%)	IA (%)
ZeroGPT	37,48%	62,52
GPTZero	51,00	49,00
Originality	0,5	99,5
Média Geral das ferramentas:	29,66	70,34

Fonte: Elaborado pelos autores

Esta tabela consolida o desempenho médio dos três verificadores na análise dos quatro textos híbridos. O ZeroGPT classificou 37,48% como humano e 62,52% como IA. O GPTZero apresentou resultado equilibrado, atribuindo 51% a humano e 49% a IA, sendo a ferramenta que melhor refletiu a natureza mista dos textos através de percentuais próximos entre as categorias. O Originality.AI classificou 99,5% como IA e apenas 0,5% como humano. A média geral das três ferramentas resultou em 70,34% de IA e 29,66% humano, com o Originality.AI impactando significativamente essa média devido à sua classificação quase categórica.

Tabela 30 - Análise por Ferramenta: ZeroGPT

ID Texto	IA (%)	Humano (%)	Média da precisão por categoria
H1 / Crítica	33,41	66,59	
H2 / Insólita	46,51	53,49	
H3 / Factual	67,24	32,76	39,747%
H4 / Dados	92,85	7,15	
HYB1 / Crítica	43,47	56,53	
HYB2 / Insólita	54,17	45,83	37,48% de Humano 62,52% de IA
HYB3 / Factual	62,21	37,79	
HYB4 / Dados	90,26	9,74	
IA1/ Crítica	100,00	0,00	
IA2 / Insólita	78,55	21,45	
IA3 / Factual	89,37	10,63	91,645%
IA4 / Dados	98,66	1,34	

Fonte: Elaborado pelos autores

Esta tabela apresenta o desempenho consolidado do ZeroGPT nos 12 textos analisados. Para os textos 100% humanos (H1-H4), a ferramenta alcançou taxa de precisão de apenas 39,747%, com resultados variados: H2 apresentou o melhor desempenho (53,49% humano), enquanto H4 obteve o pior resultado (apenas 7,15% humano). Nos textos híbridos (HYB1-HYB4), o ZeroGPT classificou 37,48% como humano e 62,52% como IA, demonstrando tendência a identificar predominância de conteúdo artificial nos textos mistos, com variação acentuada entre HYB2 (equilíbrio próximo a 50%) e HYB4 (90,26% IA). Para os textos 100% IA (IA1-IA4), a ferramenta atingiu taxa de precisão de 91,645%, com destaque para IA1 que obteve identificação perfeita (100% IA).

Tabela 31 - Análise por Ferramenta: GPTZero

ID Texto	IA (%)	Humano (%)	Híbrido (%)	Média da precisão por categoria
H1 / Crítica	0,00	96,00	4,00	
H2 / Insólita	0,00	100,00	0,00	
H3 / Factual	0,00	100,00	0,00	96,25%
H4 / Dados	4,00	89,00	7,00	
HYB1 / Crítica	83,00	12,00	5,00	
HYB2 / Insólita	7,00	92,00	1,00	51,00% de Humano 7,25% de Híbrido
HYB3 / Factual	39,00	57,00	4,00	41,75% de IA
HYB4 / Dados	35,00	46,00	19,00	
IA1/ Crítica	82,00	6,00	12,00	
IA2 / Insólita	79,00	18,00	3,00	
IA3 / Factual	54,00	0,00	46,00	91,75%
IA4 / Dados	83,00	9,00	8,00	

Fonte: Elaborado pelos autores

Esta tabela apresenta o desempenho consolidado do GPTZero nos 12 textos analisados. Para os textos 100% humanos (H1-H4), a ferramenta alcançou taxa de precisão de 96,25%, com destaque para H2 e H3 que obtiveram identificação perfeita (100% humano). Nos textos híbridos (HYB1-HYB4), o GPTZero apresentou resultados fragmentados: 51% classificados como humano, 41,75% como IA e apenas 7,25% como híbrido. Para os textos 100% IA (IA1-IA4), a ferramenta atingiu taxa de precisão de 91,75%, com destaque para IA3 que apresentou 46% de classificação híbrida, o maior percentual nessa categoria em toda a análise.

Tabela 32 - Análise por Ferramenta: Originality

ID Texto	IA (%)	Humano (%)	Média da precisão por categoria
H1 / Crítica	100,00	0,00	
H2 / Insólita	0,00	100,00	50%
H3 / Factual	0,00	100,00	
H4 / Dados	100,00	0,00	
HYB1 / Crítica	100,00	0,00	
HYB2 / Insólita	99,00	1,00	0,5% de Humano 99,5% de IA
HYB3 / Factual	99,00	1,00	
HYB4 / Dados	100,00	0,00	
IA1/ Crítica	100,00	0,00	
IA2 / Insólita	100,00	0,00	
IA3 / Factual	100,00	0,00	100%
IA4 / Dados	100,00	0,00	

Fonte: Elaborado pelos autores

Esta tabela apresenta o desempenho consolidado do Originality.AI nos 12 textos analisados. Para os textos 100% humanos (H1-H4), a ferramenta alcançou taxa de precisão de apenas 50%, com resultados polarizados: H2 e H3 foram corretamente identificados como 100% humanos, enquanto H1 e H4 foram erroneamente classificados como 100% IA. Nos textos híbridos (HYB1-HYB4), o Originality.AI apresentou desempenho extremamente baixo, classificando 99,5% como IA e apenas 0,5% como humano, desconsiderando quase totalmente a contribuição humana nos textos mistos. Para os textos 100% IA (IA1-IA4), a ferramenta atingiu taxa de precisão perfeita de 100%, identificando corretamente todos os textos como integralmente artificiais.

Tabela 33 - Comparativo entre as Ferramentas de detecção

Funções	GPTzero	ZeroGPT	Originality
Possui plano Gratuito com suporte ao idioma português-Br	X	X	
Identifica texto híbridos	X		
Destaque na frase produzida por IA	X	X	X
Reconhece conteúdo de IA proveniente do Chat GPT, GPT 3, GPT 4, GPT5	X	X	X
Funcionade de humanizador	X	X	X
Disponibiliza suporte completo ao português-Br	X	X	X
Garante taxa inferior de 2% em Falso negativo	X	X	X
Promete taxa de precisão 99%	X		X

Fonte : Elaborada pelos autores

Com base nos resultados observados nas tabelas anteriores e confrontando com as funcionalidades prometidas na Tabela 12, observam-se discrepâncias significativas entre o que as ferramentas anunciam e seu desempenho efetivo.

Enxergando os dados de forma analítica, alguns resultados apresentaram altas taxas de falsos positivos o que tende a ser preocupante para uso didático ou profissional, na tabela 34 foi possível comparar a precisão, Taxa de falso positivo e Taxa de falso negativo geral de cada verificador de IA.

Tabela 34 - Comparativo dos resultados gerais

Ferramenta	Precisão	Taxa de falso positivo	Taxa de falso negativo
ZeroGPT	33,33%	100%	0%
GPTZero	75,00%	37,50%	0%
Originality	50,00%	75,00%	0%

Fonte : Elaborado pelos autores

Apesar do desempenho relevante do verificador GPTZero, que apresentou precisão de 75% conforme evidenciado nas análises realizadas, observa-se que o resultado não corresponde integralmente às expectativas criadas por seu portfólio e materiais promocionais. Ainda que eficiente, a precisão registrada ficou restrita a 75%. É importante destacar que cada ferramenta apresentou melhor desempenho em determinados tipos de texto. O GPTZero demonstrou elevada competência na identificação dos textos produzidos por humanos e foi o único verificador capaz de reconhecer corretamente o texto híbrido como sendo de autoria humana, considerando que este havia recebido apenas correções gramaticais e ajustes de estilo.

O Originality, embora tenha identificado corretamente dois dos quatro textos integralmente humanos, evidenciou maior dificuldade na detecção de textos humanos e híbridos, apresentando menor sensibilidade nesse aspecto quando comparado ao GPTZero.

Por fim, o ZeroGPT obteve desempenho inferior em relação aos demais verificadores, indicando percentual de autoria por IA até mesmo em textos produzidos integralmente por humanos. Esse resultado reforça suas limitações na discriminação entre conteúdos autênticos e gerados por inteligência artificial.

Em síntese, embora todas as ferramentas tenham sido capazes de identificar corretamente os textos produzidos integralmente por inteligência artificial, nenhuma delas atingiu plenamente os níveis de precisão anunciados. Conforme demonstrado nas tabelas 30 a 32, foi possível notar um maior percentual de detecção de IA em textos humanos com maior complexidade, o que indica dificuldade dos detectores em lidar com conteúdos mais densos e tecnicamente estruturados.

Durante as análises o GPTZero apresentou o melhor desempenho geral, evidenciando maior consistência entre os cenários analisados. O ZeroGPT obteve resultados intermediários, com limitações especialmente notáveis na identificação de

textos humanos. Já o Originality.AI, apesar de ser uma ferramenta paga e de demonstrar excelente desempenho na detecção de textos totalmente gerados por IA, apresentou as maiores restrições nas demais categorias avaliadas.

5. CONSIDERAÇÕES FINAIS

O presente estudo avaliou o nível de confiabilidade das ferramentas comerciais de detecção de conteúdo gerado por Inteligência Artificial (ZeroGPT, GPTZero e Originality.AI), na identificação de textos jornalísticos em língua portuguesa, contemplando produções humanas, artificiais e híbridas. Os resultados demonstraram que, embora os verificadores analisados apresentem capacidade satisfatória para reconhecer textos totalmente produzidos por IA, seu desempenho é consideravelmente inferior quando expostos a textos humanos mais complexos ou conteúdos híbridos. Assim, conclui-se que o nível de confiabilidade dessas ferramentas, no escopo investigado, é limitado e insuficiente para uso como critério único de verificação de autenticidade textual.

A análise revelou que a categoria de textos híbridos representa o maior desafio para os verificadores avaliados. As ferramentas demonstraram dificuldade significativa em detectar corretamente conteúdos híbridos, que contêm interação entre autor humano e IA, tendendo a classificá-los principalmente como artificiais. Esse comportamento indica que os modelos atuais ainda não conseguem reconhecer com clareza nuances de estilo, variações linguísticas naturais e padrões de revisão humana que coexistem com trechos gerados automaticamente.

Ao contrastar os resultados obtidos com as especificações técnicas e promessas comerciais divulgadas pelas plataformas, observou-se uma discrepância relevante entre o desempenho anunciado e o desempenho real. Embora algumas ferramentas afirmem alcançar precisão próxima de 99% e taxas mínimas de falsos positivos, os experimentos demonstraram índices significativamente inferiores, especialmente na detecção de textos humanos e híbridos. Tal diferença reforça a necessidade de cautela

no uso dessas soluções e de maior transparência por parte das empresas responsáveis.

O desenvolvimento deste trabalho apresentou desafios relevantes, em especial pela escassez de referências e artigos referenciando o tema e diversas elaborações e análises de prompts para definir qual se encaixaria melhor para atender todos os parâmetros definidos. Ainda assim, foi possível alcançar os objetivos estabelecidos, apresentar evidências empíricas e contribuir para o debate acadêmico sobre a confiabilidade e os limites dessas tecnologias.

Para pesquisas futuras, recomenda-se a ampliação do corpus analisado, incluindo diferentes gêneros textuais, níveis de complexidade e áreas do conhecimento, bem como a comparação com ferramentas adicionais. Sugere-se, também, estudos que acompanhem a evolução desses detectores ao longo do tempo, investigações sobre o impacto de diferentes estratégias de prompt na identificação de textos híbridos e análises qualitativas envolvendo docentes e estudantes para compreender as implicações acadêmicas, pedagógicas e éticas do uso desses sistemas.

Em síntese, embora os detectores de IA representem avanços tecnológicos relevantes, sua utilização ainda deve ser vista como apoio complementar, e não como instrumento definitivo de autenticação textual.

6. REFERÊNCIAS BIBLIOGRÁFICAS

ALMARIE, B.; TEIXEIRA, P. E. P.; PACHECO-BARRIOS, K.; ROSSETTI, C. A.; FREGNI, F. Editorial – The Use of Large Language Models in Science: Opportunities and Challenges. **Princ Pract Clin Res.**, v. 9, n. 1, p. 1-4, jul. 2023. DOI: 10.21801/ppcrj.2023.91.1.

ANDRADE, L. M. A escrita, uma evolução para a humanidade. **Linguagem em (Dis)curso**, v. 1, n. 1, jul./dez. 2001. Disponível em: <https://portaldeperiodicos.animaeducacao.com.br/index.php/Linguagem_Discurso/article/view/167>. Acesso em: 26 set. 2025.

ANTONELLI, A. L. Desafios de grandes modelos de linguagem generativa na reprodução de complexidade textual. **Texto Livre: Linguagem e Tecnologia**, v. 16, n. 1, p. 1-20, 2025. Disponível em: <<https://periodicos.ufmg.br/index.php/textolivre/article/view/58530>>. Acesso em: 26 set. 2025.

BERGMANN, D.; STRYKER, C. **O que é um mecanismo de atenção?**. IBM, [s.d.]. Disponível em: <<https://www.ibm.com/br-pt/think/topics/attention-mechanism>>. Acesso em: 23 set. 2025.

BLANK, I. A. What are large language models supposed to model? . **Trends in Cognitive Sciences**, v. 27, n. 11, p. 987-989, nov. 2023. DOI: 10.1016/j.tics.2023.08.002. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1364661323002024>>. Acesso em: 05 out. 2025.

BORTOLANZA, A. M. E.; COSTA, S. A. F. da. A linguagem escrita: uma história de sua pré-história na infância. **Perspectiva**, Florianópolis, v. 34, n. 3, p. 928-947, set./dez. 2016. DOI: <https://doi.org/10.5007/2175-795X.2016v34n3p928>. Disponível em: <<https://periodicos.ufsc.br/index.php/perspectiva/article/view/2175-795X.2016v34n3p928>>. Acesso em: 26 set. 2025.

BRASIL. **Lei nº 9.610, de 19 de fevereiro de 1998**. Altera, atualiza e consolida a legislação sobre direitos autorais e dá outras providências. Disponível em: <https://www.planalto.gov.br/ccivil_03/Leis/L9610.htm>. Acesso em: 26 set. 2025.

BRASIL supera expectativas na maior festa esportiva do planeta. **G1**, Rio de Janeiro, dez. 2016. Disponível em: <<https://g1.globo.com/retrospectiva/2016/noticia/2016/12/brasil-supera-expectativas-na-maior-festa-esportiva-do-planeta.html>>. Acesso em: 18 out. 2025.

BRETAG, T. Contract cheating and assessment design: exploring the relationship. **Assessment & Evaluation in Higher Education**, v. 44, n. 5, p. 676–683, 2019. Disponível em:

<https://www.researchgate.net/publication/329692598_Contract_cheating_and_assessment_design_exploring_the_relationship>. Acesso em: 26 set. 2025.

CABRAS surpreendem cientistas ao identificar expressões faciais humanas. **Jornal da USP**, São Paulo, 29 ago. 2018. Disponível em: <<https://jornal.usp.br/ciencias/ciencias-biologicas/cabras-surpreendem-cientistas-ao-rec онhecer-expresso-faciais-humanas/>>. Acesso em: 18 out. 2025.

CHARTIER, R. **A aventura do livro: do leitor ao navegador**. São Paulo: Editora UNESP, 1998. Disponível em: <<https://periodicos.ufba.br/index.php/entreideias/article/view/2927>>. Acesso em: 26 set. 2025.

COPE. **Authorship**. 2019. Committee on Publication Ethics. Disponível em: <https://publicationethics.org/guidance/discussion-document/authorship?check_logged_in=1>. Acesso em: 26 set. 2025.

COPE. **Authorship and AI tools**. 2023. Committee on Publication Ethics. Disponível em: <<https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>>. Acesso em: 26 set. 2025.

DICIO. **Híbrido**. [s.d.]. Disponível em: <<https://www.dicio.com.br/hibrido/>>. Acesso em: 26 set. 2025.

ECO, U. **Como se faz uma tese**. 13. ed. São Paulo: Perspectiva, 1989. Disponível em: <https://www.academia.edu/128404344/COMO_SE_FAZ_UMA_TESE_por_Umberto_Eco_>. Acesso em: 26 set. 2025.

FORD, M. **Architects of Intelligence: The truth about AI from the people building it**. Birmingham: Packt Publishing, 2018.

FOUCAULT, M. **O que é um autor?** [1969]. Disponível em: <https://www.academia.edu/6204519/Foucault_Michel_O_que_e_um_autor>. Acesso em: 26 set. 2025.

GARCIA, A. C. B. Ética e Inteligência Artificial. **Computação Brasil**, n. 43, p. 14-22, nov. 2020. DOI: <https://doi.org/10.5753/compbr.2020.43.1791>. Disponível em: <<https://journals-sol.sbc.org.br/index.php/comp-br/article/view/1791>>. Acesso em: 17 set. 2025.

GONZALEZ, M.; LIMA, V. L. S. de. Recuperação de Informação e Processamento da Linguagem Natural. In: **CONGRESSO DA SOCIEDADE BRASILEIRA DE COMPUTAÇÃO - JORNADA DE ATUALIZAÇÃO EM INFORMÁTICA**, 23., 2003, Porto Alegre. Anais... Porto Alegre: SBC, 2003. v. 3, p. 347-395.

GOODFELLOW, I. et al. Generative adversarial networks. **Communications of the ACM**, v. 63, n. 11, p. 139-144, out. 2020. DOI: <https://doi.org/10.1145/3422622>. Disponível em: <<https://dl.acm.org/doi/10.1145/3422622>>. Acesso em: 17 set. 2025.

HEXHAM, I. **On Plagiarism and Integrity**. Humanities Computing Unit, University of Alberta, 1992. Disponível em: <https://www.researchgate.net/publication/236899249_The_Plague_of_Plagiarism_Academic_Plagiarism_Defined_Originally_published_as_On_Plagiarism_and_Integrity>. Acesso em: 26 set. 2025.

HIRSCHBERG, J.; MANNING, C. D. Advances in natural language processing. **Science**, v. 349, n. 6245, p. 261-266, jul. 2015. Disponível em: <<https://cs224d.stanford.edu/papers/advances.pdf>> Acesso em: Acesso em: 17 set. 2025.

ICMJE. **Recomendações para elaboração, redação, edição e publicação de trabalhos acadêmicos em periódicos médicos**. 2014. International Committee of Medical Journal Editors. Disponível em: <<https://www.icmje.org/recommendations/translations/portugese2014.pdf>>. Acesso em: 26 set. 2025.

IPPOLITO, D. et al. Automatic Detection of Generated Text is Easiest when Humans are Fooled. In: **ANNUAL MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS**, 58., 2020. Proceedings Association for Computational Linguistics, 2020. p. 1808–1822. Disponível em: <<https://aclanthology.org/2020.acl-main.164/>>. Acesso em: 26 set. 2025.

JESUEGRACILIANO. Impacto da IA na produção de textos acadêmicos: desafios do presente. **Jesuegraciliano (blog)**, 29 mar. 2024. Disponível em: <https://jesuegraciliano.wordpress.com/2024/03/29/impacto-da-ia-na-producao-de-textos-academicos-desafios-do-presente/?utm_source=chatgpt.com>. Acesso em: 27 set. 2025.

KAUFMAN, D.; SANTAELLA, L. O papel dos algoritmos de inteligência artificial nas redes sociais. **Revista FAMECOS**, Porto Alegre, v. 27, n. 1, p. 1-10, maio 2020. DOI: <https://doi.org/10.15448/1980-3729.2020.1.34074>. Disponível em: <<https://revistaseletronicas.pucrs.br/revistafamecos/article/view/34074>>. Acesso em: 17 set. 2025.

MATARAZZO, A.; TORLONE, R. **A Survey on Large Language Models with some Insights on their Capabilities and Limitations** [preprint]. [S. l.]: arXiv, 2025. DOI: 10.48550/arXiv.2501.04040. Disponível em: <<https://doi.org/10.48550/arXiv.2501.04040>>. Acesso em: 04 out. 2025.

MEYER, J. G. et al. ChatGPT and large language models in academia: opportunities and challenges. **BioData Mining**, v. 16, n. 20, 2023. DOI:

10.1186/s13040-023-00339-9. Disponível em:
 <<https://biodatamining.biomedcentral.com/articles/10.1186/s13040-023-00339-9>>. Acesso em: 09 out. 2025.

MINAEE, S. et al. Large Language Models: A Survey [preprint]. [S. I.]: arXiv, 2024. Disponível em: <<https://arxiv.org/abs/2402.06196>>. Acesso em: 2 out. 2025.

NOGARO, A.; PORTO, A. P. T.; PORTO, L. T. A produção escrita e a formação de professores. **Educação**, Santa Maria, v. 44, p. 1-25, 2019. DOI: <https://doi.org/10.5902/1984644430978>. Disponível em: <<https://periodicos.ufsm.br/reveducacao/article/view/30978>>. Acesso em: 26 set. 2025.

OPENAI. Apresentamos o GPT-5. **OpenAI**, 7 ago. 2025. Disponível em: <https://openai.com/pt-BR/index/introducing-gpt-5/?utm_source=chatgpt.com>. Acesso em: 04 out. 2025.

OPENAI. O GPT-4 é o sistema mais avançado da OpenAI, projetado para gerar respostas mais úteis e seguras. **OpenAI**, [2025?]. Disponível em: <<https://openai.com/pt-BR/index/gpt-4/>>. Acesso em: 04 out. 2025.

ORIGINALITY.AI. AI Content Detection & Plagiarism Checker. 2023. Disponível em: <<https://originality.ai>>. Acesso em: 26 set. 2025.

PAES, G. S. M. de C. **Detecção de textos gerados por LLMs em português**. 2025. 49 f. Monografia (Graduação em Ciência da Computação) – Instituto de Ciências Exatas e Biológicas, Universidade Federal de Ouro Preto, Ouro Preto, 2025. Disponível em: <<http://www.monografias.ufop.br/handle/35400000/7640>>. Acesso em: 20 set. 2025.

PENTZ, E. CrossRef: The Missing Link. **College & Research Libraries News**, v. 62, n. 6, p. 620–623, 2001. Disponível em: <<https://crln.acrl.org/index.php/crlnews/article/view/22135/0>>. Acesso em: 26 set. 2025.

PIMENTEL, M.; AZEVEDO, V.; CARVALHO, F. ChatGPT: a era da autoria híbrida humana/o-IA. **SBC Horizontes**, 21 mar. 2023. ISSN 2175-9235. Disponível em: <<https://horizontes.sbc.org.br/index.php/2023/03/chatgpt-a-era-da-autoria-hibrida/>>. Acesso em: 26 out. 2025.

PRIVACIDADE digital: impactos e efeitos colaterais. **Correio Braziliense**, Brasília, 17 jan. 2020. Disponível em: <https://www.correobraziliense.com.br/app/noticia/opiniao/2020/01/17/internas_opiniao%2C821038/artigo-privacidade-digital-impactos-e-efeitos-colaterais.shtml>. Acesso em: 18 out. 2025.

PROJEÇÃO da População 2018: número de habitantes do país deve parar de crescer em 2047. **Agência de Notícias IBGE**, Brasília, 25 jul. 2018. Disponível em:

<<https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/21837-projecao-da-populacao-2018-numero-de-habitantes-do-pais-deve-parar-de-crescer-em-2047>>. Acesso em: 18 out. 2025.

SCHOPOW, N.; OSTERHOFF, G.; BAUR, D. Applications of the Natural Language Processing Tool ChatGPT in Clinical Practice: Comparative Study and Augmented Systematic Review. **JMIR Medical Informatics**, v. 11, e48933, nov. 2023. DOI: 10.2196/48933. Disponível em: <<https://medinform.jmir.org/2023/1/e48933>>. Acesso em: 20 set. 2025.

SHASTRI, Y. Mecanismo de atenção em LLMs: Uma explicação intuitiva. **DataCamp**, jan. 2025. Disponível em: <<https://www.datacamp.com/pt/blog/attention-mechanism-in-langs-intuition>>. Acesso em: 23 set. 2025.

SILVA, A. P. da; FREESZ, L. Inspiração e concatenação: rascunhos do processo criativo de A hora da estrela. **Verbo de Minas**, Juiz de Fora, v. 18, n. 31, p. 23-38, jan./jul. 2017. Disponível em: <<https://seer.uniacademia.edu.br/index.php/verboDeMinas/article/download/1170/837>>. Acesso em: 26 set. 2025.

SILVA, C. M. V. da; KRÜGER, A. C. **Transcrição e análise paleográfica de manuscritos do século XIX: registros da escravidão negra em Santa Catarina**. 2021. 31 f. Trabalho de Conclusão de Curso (Graduação em Arquivologia) - Universidade Federal de Santa Catarina, Florianópolis, 2021. Disponível em: <https://repositorio.ufsc.br/bitstream/handle/123456789/235714/2021.2_CI%C3%A1udi-a-Maria-Vieira-da-Silva_TCC.pdf?sequence=1>. Acesso em: 26 set. 2025.

SINGH, S. Estatísticas mais recentes dos usuários do ChatGPT 2025 (Relatório de crescimento e uso). **DemandSage**, 7 out. 2025. Disponível em: <<https://www.demandsage.com/chatgpt-statistics>>. Acesso em: 04 out. 2025.

SWAN, M. **Blockchain: Blueprint for a New Economy**. Sebastopol: O'Reilly Media, 2015. Disponível em: <https://www.academia.edu/44112222/Melanie_Swan_Blockchain_BLUEPRINT_FOR_A_NEW_ECONOMY>. Acesso em: 26 set. 2025.

TIAN. **GPTZero – AI Detector for ChatGPT, GPT-4, GPT-3 & More**. 2023. Disponível em: <<https://gptzero.me>>. Acesso em: 26 set. 2025.

TRINDADE, A. S. C. E.; OLIVEIRA, H. P. C. Inteligência artificial (IA) generativa e competência em informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 28, p. 1-22, 2023. Disponível em: <<https://periodicos.ufmg.br/index.php/pci/article/view/47485>>. Acesso em: 26 set. 2025.

TURNITIN. A escrita por Inteligência Artificial: os atuais desafios e oportunidades para a Educação. **Turnitin Brasil**, 11 jan. 2023a. Disponível em: <<https://www.turnitin.com.br/blog/escrita-por-ia-atais-desafios-e-oportunidades-para-educacao>>. Acesso em: 27 set. 2025.

TURNITIN. Texto gerado por Inteligência Artificial (IA): a ameaça, a responsabilidade e a promessa. **Turnitin Brasil (blog)**, 18 jan. 2023b. Disponível em: <https://www.turnitin.com.br/blog/texto-gerado-por-ia-ameaca-responsabilidade-promessa?utm_source=chatgpt.com>. Acesso em: 27 set. 2025.

VASWANI, A. et al. Attention is All You Need. In: CONFERENCE ON NEURAL INFORMATION PROCESSING SYSTEMS, 31., 2017, Long Beach. **Advances in Neural Information Processing Systems 30**. Long Beach: NIPS, 2017. p.1-11. DOI: 10.48550/arXiv.1706.03762.

ZEROGPT. **AI Detector & ChatGPT Plagiarism Checker**. 2023. Disponível em: <<https://www.zerogpt.com>>. Acesso em: 26 set. 2025.

APÊNDICE A – PROMPT UTILIZADO PARA GERAÇÃO DE TEXTO COM IA

Atue como um redator profissional de jornal especializado em política e economia brasileira.

Pesquise e redija uma notícia completa e detalhada sobre o seguinte tema:

[Assunto principal], assegura [principais autoridade do assunto se tiver] , [houve auxílio de : principais atores/instituições] , ocorrido em [data ou ano ocorrido], analisando { seus impactos/desdobramentos principais}."

Instruções para o texto:

Produza uma notícia completa e atualizada, com cerca de 500 palavras, pronta para publicação (sem rótulos como 'título:', 'Subtítulo:' ou 'contexto:')." contendo :

- Título principal (chamativo, informativo e coerente com o tema).
- Subtítulo (linha fina) com breve contextualização.
- Lead (1º parágrafo) respondendo claramente: o quê, quem, quando, onde, como e por quê.
- Corpo da notícia dividido em parágrafos coesos e bem encadeados, apresentando:
- Contexto do tema;
- Repercussão do tema;
- Citações de fontes (como Polícia Rodoviária Federal, Justiça Federal, Ministério Público,sindicatos ou outros jornais);
- Impactos econômicos e sociais.
- Conclusão com análise ou desdobramentos futuros.

Use linguagem jornalística imparcial, clara e objetiva, sem adjetivos opinativos.Mencione fontes oficiais ou referências conhecidas (ex: segundo a Polícia Rodoviária Federal, de acordo com decisão da Justiça Federal em Curitiba, etc.).

O texto deve soar natural, humano e profissional, como se fosse publicado em um grande portal de notícias (exemplo: G1, Estadão, Folha de S.Paulo), Evite copiar trechos literais — reescreva com base em informações públicas.

Gere o texto completo, pronto para publicação.

APÊNDICE B – PROMPT UTILIZADO PARA GERAÇÃO DE TEXTO HÍBRIDO

Atue como um editor de notícias experiente e refinado, especializado em política e economia brasileira.

Você receberá um texto jornalístico. Sua tarefa é aprimorá-lo para publicação, garantindo que ele seja acessível, envolvente e de fácil compreensão para um público amplo, ao mesmo tempo em que mantém a precisão dos fatos e a integridade do conteúdo original.

Instruções para o aprimoramento:

1. Fluidez e Clareza Narrativa: Reorganize a ordem dos fatos, se necessário, para garantir a máxima fluidez e uma sequência lógica que torne a leitura fácil e comprehensível. O texto deve progredir de forma natural, guiando o leitor através das informações sem obstáculos.

2. Vocabulário e Acessibilidade: Avalie o vocabulário utilizado e substitua termos menos comuns, jargões ou complexos por sinônimos mais acessíveis e diretos, sem comprometer o sentido ou o tom profissional da notícia. O objetivo é que o texto seja facilmente digerível por um leitor sem conhecimento prévio aprofundado no tema.

3. Manutenção de Citações e Fontes: Todas as citações de fontes, declarações e referências presentes no texto original devem ser mantidas exatamente como estão, sem quaisquer alterações em seu conteúdo ou atribuição.

4. Estilo Jornalístico Profissional: Assegure que o texto final preserve um tom jornalístico imparcial, claro e objetivo. O aprimoramento deve resultar em um texto que seja polido, profissional e natural, como se tivesse passado por uma revisão editorial de um grande portal de notícias (exemplo: G1, Estadão, Folha de S.Paulo).

5. Polimento Geral: Revise a gramática, sintaxe e pontuação para eliminar quaisquer imperfeições, garantindo que o texto esteja impecável para publicação.

Texto a ser aprimorado: [INSIRA O TEXTO ORIGINAL AQUI]

APÊNDICE C – IA1 / Crítica

Privacidade digital e poder global: a disputa por dados que impulsionou a criação da LGPD no Brasil

Inspirada em escândalos internacionais e na corrida tecnológica entre potências, a Lei Geral de Proteção de Dados redefine a relação entre empresas, Estado e cidadãos desde 2020.

A entrada em vigor da Lei Geral de Proteção de Dados (LGPD), em agosto de 2020, marcou um ponto de virada na política digital brasileira. Criada para proteger a privacidade e garantir o controle dos cidadãos sobre suas informações pessoais, a legislação surgiu em meio a um intenso debate global sobre vigilância e uso político de dados — tema que ganhou destaque após os escândalos de Edward Snowden, em 2013, e da Cambridge Analytica, em 2016. A norma coloca o Brasil entre as nações que buscam equilibrar inovação tecnológica e soberania digital em um cenário dominado pela disputa entre China e Estados Unidos.

O caso Snowden revelou a extensão dos programas de espionagem digital conduzidos pela Agência de Segurança Nacional dos EUA (NSA), que monitoravam comunicações de cidadãos e governos em diversos países, incluindo o Brasil. Já o escândalo da Cambridge Analytica expôs o uso indevido de dados de milhões de usuários do Facebook para influenciar eleições nos Estados Unidos e no Reino Unido, em 2016. Esses episódios despertaram preocupações sobre o poder de corporações e governos na manipulação de informações pessoais e no impacto direto sobre democracias.

No Brasil, o debate sobre privacidade digital se intensificou entre 2018 e 2020, durante a tramitação da LGPD no Congresso Nacional. Inspirada no Regulamento Geral sobre a Proteção de Dados (GDPR), da União Europeia, a lei brasileira foi sancionada após intensa pressão de especialistas em cibersegurança, juristas e organizações da sociedade civil. “A LGPD é uma resposta à necessidade de proteger a soberania informacional do país e garantir transparência no uso dos dados”, afirmou à época Danilo Doneda, um dos redatores do anteprojeto da lei e referência nacional no tema.

Com a nova legislação, empresas públicas e privadas passaram a ser obrigadas a adotar medidas de segurança e transparência na coleta e tratamento de informações pessoais. O descumprimento pode resultar em multas de até 2% do faturamento anual, limitadas a R\$ 50 milhões por infração, conforme a Autoridade Nacional de Proteção de Dados (ANPD). Desde então, o setor privado tem investido pesado em consultorias especializadas, impulsionando uma nova economia de serviços voltados à conformidade regulatória e à cibersegurança.

Segundo levantamento da Brasscom (Associação Brasileira das Empresas de Tecnologia da Informação e Comunicação), o mercado de consultoria em proteção de dados movimentou mais de R\$ 2,5 bilhões entre 2020 e 2024, com forte demanda de bancos, e-commerce e órgãos públicos. “A adequação à LGPD tornou-se fator estratégico para competitividade e credibilidade das empresas”, declarou em nota a Federação das Indústrias do Estado de São Paulo (Fiesp).

Além dos impactos econômicos, a lei também tem implicações diretas sobre a democracia. Especialistas alertam que o controle sobre o fluxo de dados é hoje um instrumento de poder geopolítico, influenciado pela disputa entre o modelo estatal de vigilância da China e o corporativo, liderado por empresas dos EUA. Nesse contexto, o Brasil tenta afirmar sua autonomia digital e proteger a integridade de seus processos eleitorais, cada vez mais dependentes de plataformas digitais.

Para o Ministério Públíco Federal, a consolidação da LGPD é essencial para preservar o espaço democrático e coibir abusos. Em nota publicada em 2024, a instituição destacou que “a proteção de dados é pilar da liberdade de expressão e da confiança pública nas instituições”.

Quase cinco anos após sua entrada em vigor, a LGPD ainda enfrenta desafios na fiscalização e na educação digital da população. Mas especialistas apontam que a lei representa um avanço significativo na construção de um ecossistema tecnológico mais ético e soberano — e um passo decisivo para que o Brasil ocupe posição estratégica na nova ordem global da informação.

APÊNDICE D – IA2 / Insólita

Cabras reconhecem expressões faciais humanas e preferem rostos felizes, revela estudo internacional

Pesquisadores investigam a origem evolutiva da empatia e da cognição em animais domesticados, com participação de cientistas brasileiras.

Um estudo publicado na revista *Royal Society Open Science*, em 2018, revelou que cabras são capazes de distinguir expressões faciais humanas — e demonstram preferência por rostos sorridentes. A pesquisa, conduzida por uma equipe internacional envolvendo as cientistas brasileiras Natália de Souza Albuquerque, da Universidade de São Paulo (USP), e Carine Savalli, da Universidade Federal do ABC (UFABC), foi realizada no Santuário Buttercups, localizado em Kent, no Reino Unido, e trouxe novas perspectivas sobre a cognição animal e as relações emocionais entre humanos e outras espécies.

O experimento consistiu em apresentar a cabras fotografias de rostos humanos com diferentes expressões — uma feliz e outra zangada — posicionadas lado a lado em um cercado. Segundo o artigo, os animais tenderam a se aproximar e interagir mais frequentemente com os rostos alegres. A metodologia controlada permitiu descartar outros fatores, como cor ou luminosidade das imagens, indicando que as cabras efetivamente reconheceram sinais emocionais.

De acordo com os autores, o estudo sugere que a habilidade de interpretar emoções humanas pode não ser exclusiva de animais tradicionalmente próximos ao homem, como cães ou cavalos. “As cabras não foram domesticadas para interagir socialmente com pessoas da mesma forma que os cães, mas ainda assim conseguem perceber aspectos emocionais de nossas expressões”, afirmou Natália de Souza Albuquerque, em comunicado divulgado pela Universidade de Cambridge, que também participou da pesquisa.

A descoberta despertou interesse entre especialistas em comportamento animal e neurociência. Segundo nota da *Royal Society*, o trabalho reforça a hipótese de que a capacidade de leitura emocional em espécies domesticadas pode ter raízes evolutivas mais amplas, associadas à convivência com humanos ao longo dos séculos. “Isso amplia a discussão sobre o bem-estar animal e sobre como as interações com humanos moldam a cognição das espécies”, avaliou Carine Savalli, coautora brasileira do estudo.

Entidades de proteção animal consideraram o resultado um avanço para o debate sobre sensibilidade e comunicação interespécies. O Instituto Brasileiro de Meio

Ambiente (Ibama) afirmou, em nota, que estudos como esse “contribuem para repensar práticas de manejo e tratamento em santuários e fazendas, reforçando a importância do enriquecimento ambiental e da empatia no trato com animais domésticos e de produção”.

Além do impacto científico, o trabalho tem repercussões éticas e econômicas. Setores agropecuários e organizações de bem-estar animal discutem, desde então, como o reconhecimento da cognição e das emoções em espécies de criação pode influenciar políticas públicas, legislações e padrões de certificação de produtos.

Os pesquisadores planejam expandir os experimentos para avaliar se cabras também conseguem identificar emoções em vozes humanas ou em outros estímulos sensoriais, como movimentos corporais. O objetivo é compreender até que ponto essa percepção emocional está ligada à domesticação e à evolução da comunicação entre espécies.

Para Natália de Souza Albuquerque, a descoberta reforça uma mensagem mais ampla: “Entender a cognição dos animais é também entender nossa própria história evolutiva e a complexidade das relações que estabelecemos com outras formas de vida”.

APÊNDICE E – IA3 / Factual

EUA passam a exigir dados de redes sociais de solicitantes de visto

Medida implementada pelo governo Donald Trump amplia controle de segurança, mas levanta críticas sobre privacidade e autocensura digital

O Departamento de Estado dos Estados Unidos passou a exigir, desde 2019, que solicitantes de visto informem nomes de usuário utilizados em redes sociais nos últimos cinco anos. A medida, proposta originalmente em março de 2018 durante o governo Donald Trump, busca reforçar o controle de segurança sobre viajantes estrangeiros — mas tem gerado preocupação entre defensores da privacidade digital e organizações de direitos civis.

A exigência se aplica à maioria dos solicitantes de visto — incluindo categorias de turismo, estudo e trabalho — e requer que os candidatos informem também endereços de e-mail e números de telefone utilizados no mesmo período. Segundo o Departamento de Estado, o objetivo é “melhorar os processos de verificação de antecedentes e proteger a segurança nacional”. Estima-se que cerca de 15 milhões de pessoas por ano sejam afetadas pela nova política.

De acordo com a Casa Branca à época, a decisão segue recomendações de segurança elaboradas após atentados terroristas ocorridos na década anterior e pretende identificar potenciais riscos por meio da análise de atividades online. No entanto, organizações como a União Americana pelas Liberdades Civis (ACLU) alertaram para os riscos de invasão de privacidade e de uso discriminatório dos dados coletados. “A vigilância em massa das redes sociais cria um ambiente de autocensura e afeta desproporcionalmente comunidades imigrantes e minorias religiosas”, afirmou Hina Shamsi, diretora do Projeto de Segurança Nacional da ACLU, em comunicado divulgado na época.

A medida também despertou preocupação entre especialistas em tecnologia e privacidade digital. Analistas apontam que o monitoramento de perfis online pode levar à avaliação subjetiva de conteúdos e opiniões, prejudicando candidatos inocentes por mal-entendidos ou postagens interpretadas fora de contexto. Segundo o jornal *The Washington Post*, funcionários consulares passaram a receber treinamentos específicos para analisar o comportamento online dos solicitantes, mas críticos argumentam que não há critérios objetivos claros para essa avaliação.

Para os solicitantes, o fornecimento de informações falsas ou omissões intencionais pode acarretar graves consequências. O Departamento de Estado informou que inconsistências detectadas podem resultar na negação do visto e, em alguns casos, na proibição permanente de entrada no país. A embaixada americana no Brasil reforçou que “as informações são verificadas com bancos de dados do governo e que a transparência é essencial para o processo de aprovação”.

Economistas e consultores do setor de turismo preveem que o aumento das exigências possa impactar o fluxo de visitantes estrangeiros aos Estados Unidos, especialmente de países

emergentes. Segundo a Associação de Viagens dos EUA, as medidas de triagem mais rígidas já haviam contribuído para uma leve queda de 1,7% no número de solicitações de visto em 2020, antes da pandemia.

Especialistas em relações internacionais observam que a política reflete uma tendência global de vigilância digital crescente, impulsionada por preocupações com terrorismo, imigração e crimes cibernéticos. No entanto, o debate sobre privacidade e liberdade de expressão deve se intensificar nos próximos anos, à medida que governos e empresas ampliam o acesso a dados pessoais.

A expectativa é que futuras administrações norte-americanas revisem o alcance da medida, buscando equilibrar segurança e direitos civis. Enquanto isso, milhões de solicitantes continuam a preencher formulários mais extensos — cientes de que, nas redes sociais, até uma simples postagem pode se tornar um critério de avaliação diplomática.

APÊNDICE F – IA4 / Dados

Brasil envelhece e cresce menos: projeções do IBGE até 2060 apontam desafios para previdência e mercado de trabalho

Revisão de 2018 das projeções populacionais do IBGE indica que o país terá mais idosos do que jovens em quatro décadas, com impacto direto sobre a economia e as políticas públicas. Migração venezuelana em Roraima e queda da fecundidade ampliam as transformações demográficas.

O Brasil caminha para uma transição demográfica profunda. Segundo a revisão de 2018 da *Projeção de População* do Instituto Brasileiro de Geografia e Estatística (IBGE), o país atingirá seu pico populacional em 2047, com cerca de 233 milhões de habitantes, e passará a encolher gradualmente até 2060. A pesquisa, que atualiza as estimativas por estados e faixas etárias, mostra que o envelhecimento populacional e a queda da fecundidade transformarão a estrutura socioeconômica brasileira nas próximas décadas.

De acordo com o IBGE, a taxa de fecundidade nacional — atualmente em torno de 1,7 filho por mulher — deve cair para 1,66 em 2034 e atingir 1,66 até 2060, consolidando-se abaixo do nível de reposição populacional (2,1 filhos). Estados como São Paulo, Rio de Janeiro e Rio Grande do Sul já registram índices inferiores a 1,6, enquanto o Norte e o Nordeste, embora mais jovens, seguem a mesma tendência de redução gradual. Em Roraima, as projeções foram fortemente influenciadas pelo fluxo migratório venezuelano, que aumentou a população local em mais de 10% entre 2016 e 2018, segundo dados da Polícia Federal e do Alto Comissariado das Nações Unidas para Refugiados (Acnur).

A migração, concentrada na fronteira de Pacaraima, alterou a composição demográfica e pressionou serviços públicos em Boa Vista, capital de Roraima. Segundo a Secretaria de Assistência Social do estado, a entrada contínua de imigrantes jovens ajudou a equilibrar, de forma pontual, a pirâmide etária regional, mas também trouxe desafios de integração e emprego. “O fenômeno migratório tem um papel local relevante, mas não altera a tendência nacional de envelhecimento”, afirmou em nota o demógrafo José Eustáquio Diniz Alves, ex-pesquisador do IBGE.

A população idosa, com 65 anos ou mais, deve triplicar até 2060, passando de 9,2% para 25,5% do total. O grupo em idade ativa (15 a 64 anos), que hoje sustenta o sistema previdenciário, começará a se reduzir já na próxima década. O Ministério da Economia avalia que essa mudança exigirá novas políticas de incentivo à produtividade e à permanência de trabalhadores mais velhos no mercado. “O envelhecimento populacional representa um dos maiores desafios fiscais do país. Será necessário equilibrar a sustentabilidade da Previdência com a proteção social”, afirmou em nota técnica o Instituto de Pesquisa Econômica Aplicada (Ipea).

No mercado de trabalho, o IBGE projeta uma queda na oferta de mão de obra jovem, o que poderá elevar a demanda por automação e requalificação profissional. Setores intensivos em

força de trabalho, como agricultura, construção civil e serviços, deverão enfrentar carência de trabalhadores em regiões interioranas, especialmente no Sul e Sudeste.

Economistas ouvidos pelo *Estadão* e pela *Folha de S.Paulo* destacam que o novo perfil demográfico pode estimular políticas de imigração e de incentivo à natalidade, à semelhança do que ocorre em países europeus. Entretanto, alertam que sem investimento em educação, inovação e produtividade, o envelhecimento poderá desacelerar o crescimento econômico.

As projeções do IBGE, atualizadas anualmente, servem de base para o planejamento de políticas públicas federais e estaduais. A próxima revisão, prevista para 2026, deve incorporar os resultados do Censo Demográfico de 2022, que já confirmou a desaceleração populacional. Até lá, especialistas reforçam que o Brasil precisa se preparar para um futuro em que viverá mais — mas com menos pessoas em idade produtiva para sustentar o sistema.

APÊNDICE G – HYB 1 / Crítica

Entre Snowden e Cambridge Analytica: a guerra global pelos dados e a encruzilhada da privacidade digital

Edward Snowden foi protagonista de um dos maiores escândalos da era moderna ao revelar, em 2013, o esquema de vigilância global da Agência de Segurança Nacional dos Estados Unidos (NSA). Três anos depois, Brittany Kaiser tornou-se um dos nomes centrais do caso Cambridge Analytica, ao expor a coleta e o uso indevido de dados de milhões de usuários do Facebook na campanha que elegeu Donald Trump.

Esses dois episódios redefiniram o debate sobre privacidade digital e continuam ecoando até hoje. Snowden e Kaiser, que antes denunciaram abusos, agora figuram como referências globais na defesa da proteção de dados — tema que ganha nova relevância com a chegada da Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil.

A nova corrida do século: quem controla os dados, controla o futuro

Em um mundo cada vez mais digital, dados se tornaram um dos recursos mais valiosos do planeta — tão estratégicos quanto petróleo ou armamentos. A “próxima guerra” não será travada em campos de batalha, mas em servidores e redes, onde informações alimentam algoritmos capazes de moldar comportamentos, economias e até democracias.

Nesse embate tecnológico, as duas potências que disputam a liderança são conhecidas: Estados Unidos e China. A grande questão é se a corrida global por regulamentações de privacidade será capaz de equilibrar essa disputa — ou de incliná-la ainda mais para um dos lados.

Democracia versus controle: duas visões de mundo

De um lado, as nações ocidentais tentam criar regras que limitem o uso de dados pessoais, na esperança de preservar princípios democráticos. A proposta é simples em teoria: garantir que informações como nome, renda, localização, e-mail e hábitos online sejam protegidas por leis e tratadas com transparência.

Do outro, governos mais autoritários buscam consolidar o conceito de “soberania digital” — o controle total sobre os dados gerados por seus cidadãos. É por esse caminho que a China avança, utilizando sistemas de reconhecimento facial, monitoramento de transações financeiras e acesso a registros médicos e acadêmicos. Essa postura tem proporcionado ao país uma vantagem tecnológica expressiva frente ao Ocidente.

O paradoxo é evidente: enquanto democracias lutam para evitar abusos e proteger direitos individuais, regimes com maior liberdade para explorar dados em larga escala acumulam poder e conhecimento. Como a inteligência artificial depende diretamente do volume e da qualidade

das informações disponíveis, essa assimetria tende a ampliar a distância tecnológica entre as potências.

O império dos dados e a expansão das big techs

A mais de 10 mil quilômetros do Vale do Silício, gigantes chinesas como Alibaba, Tencent e Baidu já ultrapassaram fronteiras e conquistaram posições de destaque entre as empresas mais valiosas do mundo — lado a lado com os tradicionais nomes americanos: Amazon, Apple, Google, Microsoft e Facebook.

Esse domínio global do mercado digital tem impulsionado também um novo setor de serviços. Escritórios de advocacia, empresas de auditoria e consultorias especializadas em privacidade — muitas delas lideradas por figuras como Edward Snowden e Brittany Kaiser — se multiplicam para atender à crescente demanda por proteção de dados.

A encruzilhada brasileira

No Brasil, esse movimento ganha força com a entrada em vigor da LGPD, em agosto de 2020. Inspirada no Regulamento Geral de Proteção de Dados da União Europeia (GDPR), aprovado em 2016, a legislação brasileira busca estabelecer limites claros para a coleta, o armazenamento e o uso de informações pessoais por empresas e governos.

O modelo europeu tornou-se referência mundial — o chamado “padrão-ouro” da privacidade digital — e vem desencadeando um efeito dominó, com diversos países adotando regras semelhantes.

Ainda assim, o futuro permanece incerto. Nenhum conjunto de normas será capaz, por si só, de servir como um “GPS” moral para a era digital. Por outro lado, a concentração do poder informacional — sob o pretexto de proteger ou controlar — pode redefinir a liderança global na revolução tecnológica em curso.

Em um mundo dividido entre liberdade e vigilância, o desafio não é apenas proteger dados, mas decidir quem deve controlá-los — e a que custo.

APÊNDICE H – HYB 2 / Insólita

Cabras reconhecem expressões faciais humanas e preferem rostos felizes, aponta estudo

Um estudo recém-publicado na prestigiada revista científica Royal Society Open Science revelou um comportamento surpreendente: cabras são capazes de distinguir expressões faciais humanas de alegria e raiva — e mostram preferência pelos rostos felizes.

A pesquisa foi conduzida por um grupo internacional de cientistas e contou com a participação das brasileiras Natália de Souza Albuquerque e Carine Savalli, da Universidade Federal de São Paulo (Unifesp) e do Instituto de Psicologia (IP) da USP.

“Esse estudo traz um resultado fascinante, uma vez que mostra que a habilidade complexa de perceber as emoções humanas por meio de dicas sutis da face não está presente somente em animais domesticados para companhia, como os cães”, explica Carine Savalli, docente da Unifesp e colaboradora no programa de pós-graduação em Psicologia Experimental do IP-USP.

Animais e emoções humanas

Ainda não se sabe o quão difundida é, no reino animal, a capacidade de reconhecer expressões faciais humanas. Estudos anteriores já haviam mostrado que cães e cavalos conseguem identificar emoções humanas — algo que pode estar ligado ao processo de domesticação, que favoreceu indivíduos mais aptos a interagir com pessoas.

Mas, diferentemente desses animais, cabras e outros bichos de fazenda não foram domesticados para convivência ou cooperação com humanos. Sua criação tem fins práticos — como produção de leite, lã ou carne. Isso levantava uma questão: seriam elas, mesmo assim, capazes de interpretar nossas expressões faciais?

O experimento

A ideia do estudo surgiu quando Natália Albuquerque, então doutoranda do Instituto de Psicologia da USP, realizava parte de sua pesquisa na Universidade de Lincoln, na Inglaterra.

“Em 2016 eu estava na Inglaterra e fui convidada pelo professor Alan McElligott para dar uma palestra sobre o meu trabalho. Tínhamos publicado um artigo que mostrava que cães conseguem reconhecer expressões emocionais humanas, tanto pela face como pela voz, e o professor McElligott, que estava estudando cabras, queria começar a olhar para a questão afetiva nesses animais”, conta Natália.

Os experimentos foram realizados no Santuário Buttercups, instituição que acolhe cabras resgatadas de maus-tratos, localizada a cerca de 75 quilômetros de Londres.

Durante os testes, cada cabra era colocada diante de duas fotografias de uma pessoa desconhecida: em uma, o rosto exibia um sorriso; na outra, uma expressão de raiva.

“A gente queria ver se as cabras mostram uma tendência a se aproximar, a interagir, a olhar mais para as faces positivas ou para as negativas. Assim poderíamos concluir que elas conseguem discriminar. Se não, elas olhariam e interagiriam com as duas da mesma forma”, explica a pesquisadora.

O resultado foi claro: as cabras se mostraram mais interessadas nos rostos felizes, sugerindo que são capazes de identificar e distinguir diferentes expressões humanas. O estudo, porém, ainda não determina se os animais preferem rostos alegres ou apenas evitam os zangados.

Rigor científico

Na etologia — ciência que estuda o comportamento animal —, o controle das variáveis é essencial para garantir que as reações observadas sejam realmente causadas pelo estímulo analisado.

“Nos estudos de comportamento e cognição temos que controlar tudo. Por isso demoram tanto para serem feitos. Outras áreas publicam bem mais rápido, exatamente porque não têm que lidar com esse tipo de viés”, afirma Natália.

Cada animal participou de quatro rodadas de testes, com intervalos de duas semanas entre elas, para evitar que se acostumassem à tarefa. Os pesquisadores também alternaram o lado de exibição das imagens — direita ou esquerda — e usaram fotografias tanto de homens quanto de mulheres.

Esse controle permitiu observar um detalhe curioso: as cabras interagiram mais com rostos felizes quando estes estavam posicionados à sua direita.

A explicação pode estar no modo como o cérebro processa emoções. Segundo uma hipótese comum na etologia, o hemisfério direito do cérebro tende a lidar com emoções negativas, enquanto o esquerdo processa emoções positivas. Assim, o fato de as cabras reagirem mais a rostos felizes no lado direito sugeriria que essa resposta emocional está ligada ao hemisfério esquerdo — reforçando a hipótese.

Novas perguntas

Para Carine Savalli, os resultados abrem novas frentes de pesquisa.

“Este estudo traz a primeira evidência da presença de uma habilidade cognitiva bastante complexa em cabras, sendo importante, portanto, que mais estudos sejam conduzidos, com amostras diferentes em condições diferentes, para reforçar os achados”, afirma.

Os cientistas agora planejam novos experimentos para aprofundar o entendimento sobre essa capacidade e investigar sua origem evolutiva.

“As cabras do santuário interagem todos os dias com pessoas, e é uma interação positiva. Então fica a pergunta: será que [elas são capazes de identificar expressões humanas] porque ao longo da história de vida delas elas têm esse contato com os humanos?”, questiona Natália.

Outra possibilidade é que essa habilidade tenha surgido como subproduto de outros traços selecionados durante a domesticação, como a mansidão — ou até mesmo que a percepção de emoções humanas seja uma característica compartilhada por diferentes grupos de animais, não apenas os domésticos.

Para Natália Albuquerque, compreender o comportamento animal vai muito além da curiosidade científica.

“Nós não somos quem nós somos sem a nossa interação com os animais. Não existe nenhuma pessoa no planeta que não interaja com animais. Por isso entender essas relações é extremamente importante para nós”, conclui.

APÊNDICE I – HYB 3 / Factual

Visto americano: agora é obrigatório informar redes sociais para entrar, estudar ou trabalhar nos EUA

Quem planeja viajar para os Estados Unidos — seja a passeio, a trabalho ou para estudar — deve redobrar a atenção com o que publica nas redes sociais. O Departamento de Estado americano anunciou que, a partir deste mês, a maioria dos solicitantes de visto precisará informar os perfis que mantém ou manteve em plataformas como Facebook, Twitter e outras, nos últimos cinco anos.

Além disso, será necessário fornecer endereços de e-mail e números de telefone utilizados nesse mesmo período. As novas regras, propostas originalmente em março de 2018, atendem a uma determinação do então presidente Donald Trump e têm como objetivo, segundo o governo, reforçar a segurança nacional e aprimorar os mecanismos de triagem de solicitantes.

“Estamos trabalhando constantemente para encontrar mecanismos que melhorem nossos processos de triagem para proteger os cidadãos americanos, ao mesmo tempo em que apoiamos as viagens legais aos Estados Unidos”, afirmou o Departamento de Estado em comunicado.

Quem deve informar as redes sociais

De acordo com as novas diretrizes, praticamente todos os solicitantes de visto — inclusive turistas, estudantes e trabalhadores — devem informar os nomes de usuários que utilizam ou utilizaram nas redes sociais. Apenas requerentes de vistos diplomáticos e oficiais estarão isentos.

Antes da mudança, essa exigência era restrita a candidatos considerados de “alto risco”, como aqueles que haviam estado em regiões controladas por grupos extremistas ou oriundos de determinados países.

Os novos formulários listam uma série de plataformas populares e pedem que o solicitante informe suas contas ativas e antigas. Há ainda um campo opcional para incluir perfis em redes que não constam na lista oficial.

Fontes de imigração ouvidas pelo jornal americano The Hill alertaram que omitir ou mentir sobre o uso das redes sociais pode gerar “sérias consequências na imigração”.

Reações e críticas

Quando a proposta foi apresentada em 2018, o governo estimou que cerca de 14,7 milhões de pessoas por ano seriam afetadas. A medida, no entanto, recebeu críticas de organizações civis e usuários de redes sociais.

A União Americana pelas Liberdades Civis (ACLU) afirmou na época que “não há evidências de que tal monitoramento seja efetivo ou justo” e alertou que a exigência pode levar os solicitantes a se autocensurarem na internet — temendo que publicações ou brincadeiras sejam interpretadas de forma negativa.

Nas redes, internautas ironizaram a política, apelidando-a de “Big Brother americano”. O jornalista Peter Clarke, ex-funcionário da rede ABC na Austrália, classificou as novas regras como “ainda mais invasivas”, observando que outros destinos turísticos “passam a acenar” para viajantes que buscam alternativas. Já o jornalista indiano Kumar Manish alertou no Twitter:

“Pensem duas, três vezes antes de postarem qualquer coisa. Seu sonho americano pode ir por água abaixo com base nisso.”

APÊNDICE J – HYB 4 / Dados

População brasileira deve parar de crescer em 2047 e encolher até 2060, aponta IBGE

A população brasileira continuará crescendo até **2047**, quando deverá atingir o pico de **233,2 milhões de habitantes**. A partir daí, o país entrará em um processo gradual de redução populacional, chegando a **228,3 milhões de pessoas em 2060**, segundo a **revisão 2018 da Projeção de População do IBGE**.

O estudo detalha a evolução demográfica do Brasil até 2060, considerando as principais variáveis — **fecundidade, mortalidade e migração** — e projetando o número de habitantes em cada uma das 27 unidades da federação. A pesquisa é realizada em parceria com órgãos de planejamento estaduais e segue recomendações da **Divisão de População das Nações Unidas**.

Envelhecimento e mudança na estrutura etária

O país vive um processo de **envelhecimento acelerado**. Em **2060, um quarto dos brasileiros (25,5%) terá 65 anos ou mais**, o equivalente a **58,2 milhões de idosos**. Em 2018, essa parcela era de apenas **9,2% (19,2 milhões)**. Já os jovens de até 14 anos cairão de **21,9% (44,5 milhões)** para **14,7% (33,6 milhões)** no mesmo período.

Essa transformação afeta a **razão de dependência**, que mede a relação entre pessoas consideradas economicamente dependentes (com menos de 15 ou mais de 64 anos) e aquelas em idade ativa (de 15 a 64 anos). Em 2018, essa razão era de **44%**, mas deve subir para **67,2% em 2060**, indicando que haverá **67 pessoas dependentes para cada 100 em idade produtiva**.

O ponto de equilíbrio deverá ocorrer por volta de **2039**, quando a proporção de jovens (25,7%) e idosos (25,8%) será praticamente igual. Desde o valor mínimo registrado em **2017 (44%)**, o indicador volta a crescer de forma contínua.

Fecundidade em queda e maternidade mais tardia

A **taxa de fecundidade total**, que representa o número médio de filhos por mulher, continua em declínio. Em **2018**, o índice é de **1,77 filho por mulher**, e a projeção indica nova queda para **1,66 em 2060**.

O IBGE destaca também uma **mudança no perfil etário da maternidade**. As mulheres têm filhos cada vez mais tarde: a **idade média** passou de **27,2 anos em 2018** e deverá alcançar **28,8 anos em 2060**, refletindo o envelhecimento do padrão reprodutivo no país.

Roraima lidera taxa de fecundidade; Distrito Federal tem a menor

Atualmente, os estados do Norte — especialmente **Roraima (2,31 filhos por mulher)**, **Amazonas (2,28)**, **Acre (2,22)** e **Amapá (2,11)** — ainda apresentam taxas acima do nível de reposição populacional (2,1).

Em 2060, **Roraima continuará na liderança**, com **1,95 filho por mulher**, seguido por **Pará, Amapá, Maranhão, Mato Grosso e Mato Grosso do Sul** (todos com **1,80**). As menores taxas devem ocorrer no **Distrito Federal (1,50)** e nos estados de **Goiás, Rio de Janeiro e Minas Gerais (1,55)**.

Taxas de fecundidade total – Brasil e Unidades da Federação (2000–2060)

(número médio de filhos por mulher)

Região / UF	2000	2010	2015	2018	2020	2030	2040	2050	2060
Brasil	2,32	1,75	1,80	1,77	1,76	1,72	1,69	1,67	1,66
Norte	3,14	2,21	2,11	2,03	2,00	1,88	1,83	1,81	1,80
Rondônia	2,55	1,84	1,88	1,88	1,87	1,82	1,79	1,77	1,76
Acre	3,94	2,45	2,37	2,22	2,14	1,89	1,80	1,77	1,75
Amazonas	3,43	2,47	2,39	2,28	2,22	2,01	1,89	1,83	1,79
Roraima	3,50	2,48	2,30	2,31	2,25	2,06	1,99	1,96	1,95
Pará	3,09	2,14	1,99	1,92	1,89	1,82	1,80	1,80	1,80
Amapá	3,82	2,55	2,29	2,11	2,04	1,86	1,82	1,80	1,80
Tocantins	2,73	2,03	1,96	1,93	1,91	1,83	1,80	1,79	1,78
Nordeste	2,59	1,82	1,78	1,75	1,73	1,69	1,67	1,66	1,66
Maranhão	3,25	2,12	1,95	1,93	1,90	1,83	1,81	1,80	1,80

Piauí	2,70	1,80	1,77	1,76	1,75	1,71	1,70	1,69	1,69
Ceará	2,60	1,75	1,74	1,70	1,69	1,66	1,65	1,64	1,64
Rio Grande do Norte	2,45	1,73	1,72	1,65	1,65	1,64	1,64	1,64	1,64
Paraíba	2,45	1,76	1,79	1,76	1,74	1,69	1,67	1,65	1,64
Pernambuco	2,46	1,79	1,83	1,76	1,74	1,68	1,66	1,65	1,64
Alagoas	2,90	2,00	1,83	1,76	1,74	1,69	1,68	1,68	1,68
Sergipe	2,68	1,82	1,78	1,74	1,72	1,67	1,66	1,65	1,65
Bahia	2,37	1,73	1,68	1,69	1,68	1,64	1,63	1,62	1,62
Sudeste	2,07	1,63	1,73	1,70	1,70	1,68	1,65	1,63	1,61
Minas Gerais	2,14	1,57	1,62	1,62	1,61	1,60	1,58	1,57	1,55
Espírito Santo	2,17	1,73	1,85	1,83	1,83	1,79	1,75	1,71	1,67
Rio de Janeiro	2,03	1,59	1,75	1,74	1,73	1,68	1,64	1,59	1,55
São Paulo	2,04	1,67	1,77	1,72	1,72	1,70	1,68	1,67	1,65
Sul	2,15	1,63	1,77	1,74	1,74	1,72	1,71	1,69	1,68
Paraná	2,19	1,72	1,80	1,80	1,79	1,76	1,74	1,71	1,68
Santa Catarina	2,10	1,60	1,74	1,74	1,74	1,72	1,71	1,69	1,68
Rio Grande do Sul	2,13	1,56	1,74	1,68	1,68	1,68	1,68	1,68	1,68
Centro-Oeste	2,16	1,77	1,90	1,87	1,85	1,80	1,75	1,69	1,63
Mato Grosso do Sul	2,33	1,92	2,04	2,02	2,01	1,96	1,91	1,85	1,80
Mato Grosso	2,35	1,88	2,08	2,06	2,05	1,99	1,92	1,86	1,80
Goiás	2,02	1,69	1,82	1,79	1,77	1,72	1,66	1,61	1,55
Distrito Federal	2,13	1,65	1,71	1,68	1,67	1,63	1,59	1,54	1,50

Fonte: IBGE – Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais.

A projeção também considerou os impactos da **migração venezuelana**. Entre **2015 e 2017, 20,5 mil venezuelanos** entraram em Roraima — número que deve chegar a **79 mil até 2022**, um aumento de **185,4%**. Segundo o IBGE, embora o saldo migratório do país se mantenha próximo de zero, **Roraima é uma exceção** devido ao impacto local da crise humanitária da Venezuela.

Esperança de vida: Santa Catarina deve manter liderança até 2060

A **esperança de vida ao nascer** continuará crescendo em todo o país. **Santa Catarina**, com **79,7 anos em 2018**, deve manter a maior média nacional e alcançar **84,5 anos em 2060**. No extremo oposto, o **Piauí** deve ter a menor expectativa de vida do país (**77 anos**).

Entre os homens, os maiores valores projetados são de **81,5 anos em Santa Catarina e 80,9 no Rio Grande do Sul**; os menores, de **72,7 no Piauí e 73,6 no Pará**. Entre as mulheres, a liderança permanece com **Santa Catarina (87,6 anos)** e **Paraná (87,0)**, enquanto **Rondônia (80,3)** e **Roraima (80,8)** terão as menores médias.

O IBGE prevê uma **nova revisão das taxas de esperança de vida e mortalidade infantil em 2022**, após a realização do **Censo Demográfico 2020**.

Panorama geral

De acordo com o IBGE, o Brasil tinha **208,5 milhões de habitantes em 2018**. O número crescerá até **233,2 milhões em 2047**, e depois cairá para **228,3 milhões em 2060**, nível equivalente ao de **2034 (228,4 milhões)**.

A transição marca o início de uma nova fase demográfica no país: **menos nascimentos, maior longevidade e envelhecimento populacional acelerado**.

ANEXO A – H1 / Crítica

Artigo privacidade digital, impactos e efeitos colaterais

Edward Snowden foi protagonista de um dos maiores escândalos da era moderna, ao denunciar o esquema de vigilância global da Agência de Segurança Nacional norte-americana nos idos de 2013. Três anos depois, Brittany Kaiser foi um dos pivôs do caso da Cambridge Analytica, ao relatar a coleta e o uso de dados de usuários do Facebook na campanha que elegeu Donald Trump. Ambos os escândalos impulsionaram a polêmica sobre privacidade digital. Os dois personagens ressurgem agora como gurus da proteção de dados em um debate em escala global — que acontece às vésperas da adoção da Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil.

Dados são um dos recursos mais valiosos do planeta. Estratégicos para vencer a próxima guerra, que, em vez de ser travada nas trincheiras em campo aberto, tem como cenário o mundo digital. Esses dados alimentam algoritmos, que têm o potencial para mudar o mundo como o conhecemos hoje. As principais forças desse embate são conhecidas: China e Estados Unidos. A pergunta que se impõe nesse cenário é se a corrida pela implantação de medidas de proteção de dados pode pender a balança de um lado e para o outro.

Nesse contexto polarizado, um lado segue na tentativa de regulamentar um mundo aparentemente irrefreável, na esperança de proteger a própria democracia, com leis e normas que devem ser observadas por todos e que visam proteger nome, dados sobre renda, localização, e-mail ou qualquer outra informação que permita identificar uma pessoa.

Em direção contrária, em outros países, o governo tenta exercer o máximo controle possível, impondo o que começa a ser chamado de “soberania digital”. Esse monitoramento envolve dados de reconhecimento facial, informações médicas e acadêmicas e operações financeiras. É por onde a China avança, ganhando um corpo de vantagem.

Então, o que vemos é um claro efeito colateral desse esforço para impedir a erosão da confiança na democracia, ao oferecer uma vantagem competitiva a quem

opera em uma lógica diferente do Ocidente. Esse cenário promete não apenas acirrar essa batalha, mas especialmente desequilibrar a balança em favor da Ásia, onde alguns governos têm acesso irrestrito a qualquer tipo de dados, sem ter que prestar contas para a sociedade. Se lembrarmos que inteligência artificial tem em sua base os dados que fazem seus algoritmos funcionarem na lógica de quanto mais melhor, podemos ter uma ideia da vantagem que estão pavimentando.

A mais de 10 mil quilômetros de distância do Vale do Silício, gigantes como Alibaba (e-commerce), Tencent (social media) e Baidu (mecanismo de busca online), que dominam o mercado nacional, avançam para além da muralha, ocupando o topo do ranking da lista de empresas mais valiosas do mundo. Ao lado dos tradicionais gigantes americanos como Amazon, Apple, Google, Microsoft e Facebook.

Do lado de cá do globo, esse movimento tem alavancado todo um novo segmento de mercado, no qual advogados, auditores e toda uma gama de profissionais oferecem consultoria em privacidade, a exemplo de Edward Snowden e Brittany Kaiser.

Um fenômeno que também se nota no Brasil, onde a LGPD entra em vigor em agosto de 2020, nos moldes do Regulamento Geral sobre a Proteção de Dados da Europa (GDPR, do inglês General Data Protection Regulation), aprovado em 2016. A lei europeia se transformou no novo padrão-ouro, em um efeito dominó, com um país após o outro adotando legislações semelhantes. Impossível prever o futuro, mas o presente se parece muito com uma encruzilhada. Afinal, por mais bem-intencionada que seja, um conjunto de normas e regulamentações não funcionará como um GPS para o caminho do bem da Era Digital. E, em contrapartida, a gestão da informação online, que parece fazer-se necessária para salvaguardar certas práticas predatórias, pode implicar uma nova liderança global da revolução tecnológica.

Fonte: Correio Braziliense (2020)

ANEXO B – H2 / Isólita

Cabras surpreendem cientistas ao identificar expressões faciais humanas

Um estudo que acaba de ser publicado na prestigiada revista científica *Royal Society Open Science* sugere que cabras são capazes de discriminar entre expressões humanas alegres e de raiva, mostrando preferência pelas primeiras. O trabalho foi realizado por um grupo internacional de pesquisadores, e contou com a participação de duas cientistas brasileiras, Natália de Souza Albuquerque

“Esse estudo traz um resultado fascinante, uma vez que mostra que a habilidade complexa de perceber as emoções humanas por meio de dicas sutis da face não está presente somente em animais domesticados para companhia, como os cães”, diz Carine Savalli, que além de ser docente na Unifesp, é colaboradora no programa de pós-graduação em Psicologia Experimental do Instituto de Psicologia (IP) da USP.

Não se sabe ao certo o quanto está difundida no reino animal a habilidade para identificar expressões faciais humanas. O que já era observado é que alguns animais domésticos, como cães e cavalos, são capazes de reconhecer algumas das nossas emoções. Porém, ambos exemplos representam espécies que foram domesticadas para realizar tarefas que requerem interação com humanos, como caça ou pastoreio – e isso abre a possibilidade de que, no processo de domesticação tenham sido selecionados os indivíduos com maior aptidão para interpretar emoções humanas. Já no caso das cabras e de outros animais de fazenda, a interação com humanos é menos relevante: a criação é motivada por características físicas, como o tipo de pelo, a quantidade e qualidade do leite ou o tamanho dos ovos. Ficava então a dúvida: será que mesmo assim esses animais poderiam interpretar expressões humanas?

O Estudo

A ideia do trabalho surgiu há dois anos, quando Natália Albuquerque, doutoranda do Instituto de Psicologia (IP) da USP, desenvolvia uma parte de sua pesquisa na Universidade de Lincoln, na Inglaterra.

Em 2016 eu estava na Inglaterra e fui convidada pelo professor Alan McElligott para dar uma palestra sobre o meu trabalho. Tínhamos publicado um artigo que mostrava que cães conseguem reconhecer expressões emocionais humanas, tanto pela face como pela voz, e o professor McElligott, que estava estudando cabras, queria começar a olhar para a questão afetiva nesses animais.

Nas suas pesquisas com cabras, McElligott tem a colaboração do Santuário de Buttercups, uma instituição de caridade que acolhe cabras maltratadas ou abandonadas, e que se encontra localizado a 75 quilômetros de Londres, na Inglaterra. Foi aqui que os testes para avaliar a capacidade das cabras para reconhecer faces humanas foram desenvolvidos.

Durante cada teste, cabras e bodes foram apresentados com duas fotografias, uma do lado da outra, de uma pessoa que elas nunca tinham visto. Em uma das fotografias a pessoa sorria, na outra mantinha uma expressão de raiva.

“A gente queria ver se as cabras mostram uma tendência a se aproximar, a interagir, a olhar mais, para as faces positivas ou para as negativas. Assim poderíamos concluir que elas conseguem discriminar. Se não, elas olhariam e interagiriam com as duas das mesma forma”, explica Natália.

Quando os pesquisadores deixaram as cabras explorar as fotografias, os animais mostraram uma preferência pelos rostos felizes, o que sugere que elas são capazes de identificar e discriminar expressões faciais humanas. O que não se sabe ao certo é se elas escolhem os rostos felizes porque os preferem, ou se o que acontece é que elas evitam os rostos zangados.

Ciência rigorosa

Um aspecto crucial nos estudos em etologia, a ciência que estuda o comportamento animal – e que tem representantes respeitados como César Ades – é o de controlar ao máximo todos os parâmetros durante os experimentos. É a única

forma de se ter certeza que um comportamento está sendo causado por um determinado estímulo e não por outro.

"Nos estudos de comportamento e cognição temos que controlar tudo. Por isso demoram tanto para serem feitos. Outras áreas publicam bem mais rápido, exatamente porque não têm que lidar com esse tipo de viés", afirma Natália.

Por exemplo, cada animal repetiu o ensaio quatro vezes, mas os pesquisadores deixaram passar duas semanas entre cada um dos ensaios para evitar que as cabras se habituassem ao teste. Outro aspecto que foi controlado foi o lado do qual cada cabra foi segurada antes de deixá-la explorar as imagens. E foram utilizadas fotografias tanto de um homem como de uma mulher.

Esse controle do cenário permitiu aos pesquisadores fazer uma observação interessante: as cabras interagiram mais com as expressões felizes quando estas foram colocadas no seu lado direito.

A razão desse comportamento não está clara, mas os pesquisadores especulam que poderia ser devido à forma como o cérebro de alguns animais processa os diferentes tipos de estímulos. De acordo com uma ideia com a que se trabalha na área da etologia, o hemisfério direito do cérebro seria responsável por processar majoritariamente emoções negativas, enquanto o hemisfério esquerdo processaria emoções positivas. Essa seria a razão pela qual os cães, quando são apresentados com estímulos negativos, tendem a fugir para o lado esquerdo – o lado oposto do hemisfério cerebral que processa o estímulo. O fato de que as cabras sejam mais receptivas a imagens de rostos felizes quando elas são mostradas no lado direito sugeriria então que as emoções estão sendo processadas pelo hemisfério esquerdo, o que sustentaria a hipótese.

Os próximos passos

Para a professora Carine, os novos resultados, longe de fechar o tema, trazem consigo muitas outras perguntas.

“Este estudo traz a primeira evidência da presença de uma habilidade cognitiva bastante complexa em cabras, sendo importante, portanto, que mais estudos sejam conduzidos, com amostras diferentes em condições diferentes, para reforçar os achados”, afirma.

Os pesquisadores já estão planejando esses novos experimentos, que permitirão conhecer o alcance do achado e, quem sabe, a origem evolutiva da capacidade das cabras para identificar emoções humanas.

“As cabras do santuário interagem todos os dias com pessoas, e é uma interação positiva. Então fica a pergunta: será que [elas são capazes de *identificar expressões humanas*] porque ao longo da história de vida delas elas têm esse contato com os humanos?”, diz Natália.

Outra possibilidade seria que, embora as cabras não tenham sido domesticadas pela sua capacidade para interagir com seres humanos, essa qualidade tenha vindo junto com outras que sim foram selecionadas, como a mansidão. Ou pode ser ainda que a capacidade para reconhecer emoções humanas não esteja restrita a animais domésticos, mas seja comum a certos grupos de animais.

Para Natália Albuquerque, responder a essas perguntas e avançar no conhecimento do comportamento animal não é um simples ato de curiosidade, mas uma forma de abordar uma questões mais complexas, filosóficas, inclusive.

“Nós não somos quem nós somos sem a nossa interação com os animais. Não existe nenhuma pessoa no Planeta que não interaja com animais. Por isso entender essas relações é extremamente importante para nós”.

Fonte: Jornal da USP (2018)

ANEXO C – H3 / Factual

Visto americano: Por que você vai ter que informar dados de suas redes sociais se quiser fazer turismo, estudar ou trabalhar nos EUA

Se você pensa em pedir ou renovar um visto dos Estados Unidos é melhor analisar o que você posta no Facebook ou no Twitter, por exemplo.

É que, segundo anunciou o Departamento de Estado americano na sexta-feira, a partir deste mês, a maioria das pessoas que se candidatar a um visto no país terá que enviar junto com a solicitação informações sobre as redes sociais que usa ou usou nos últimos cinco anos.

Mas não é só isso.

De acordo com as novas regras, além de toda a documentação que já precisam apresentar, os interessados também devem fornecer endereços de e-mail e números de telefone que utilizaram também nos últimos cinco anos.

As novas exigências foram propostas pela primeira vez em março de 2018. Elas atendem a uma determinação do presidente Donald Trump, e, segundo o governo, buscam aumentar o controle de segurança sobre os solicitantes de visto.

"Estamos trabalhando constantemente para encontrar mecanismos que melhorem nossos processos de triagem para proteger os cidadãos americanos, ao mesmo tempo em que apoiamos as viagens legais aos Estados Unidos", disse o Departamento.

Quem deverá fornecer informações sobre as redes sociais?

De acordo com as novas regras, a maioria das pessoas que se candidata a um visto dos EUA deve enviar informações como os nomes de usuários que mantém ou mantinha nas redes sociais.

A obrigatoriedade vale para quem vai aos Estados Unidos fazer turismo, trabalhar ou estudar. Já certos solicitantes de vistos diplomáticos e oficiais estarão isentos das novas medidas, segundo o Departamento de Estado.

Anteriormente, apenas candidatos que demandavam uma investigação adicional por parte das autoridades - como os que haviam estado em partes do mundo controladas por grupos radicais ou eram provenientes de determinados países - tinham essa obrigação.

Os novos formulários de solicitação de visto listam uma série de redes sociais e exigem que o solicitante informe os nomes das contas que têm ativas ou as que utilizou nos últimos cinco anos.

A solicitação também inclui a opção de oferecer informações voluntariamente sobre perfis que possuam ou possuíram em outras redes, mas que não constam na relação disponível na ficha.

Segundo fontes de imigração consultadas pelo jornal americano The Hill, quem mentir sobre o uso das redes sociais poderá enfrentar "sérias consequências na imigração".

Qual foi a reação às novas medidas?

Quando a proposta de pedir informações sobre as redes sociais dos solicitantes de visto foi apresentada em 2018, as autoridades estimaram que ela afetaria 14,7 milhões de pessoas por ano.

A União Americana pelas Liberdades Civis fez críticas na época, afirmando que "não há evidências de que tal monitoramento das redes seja efetivo ou justo" e considerando que ele levaria as pessoas a se autocensurarem na internet, já que qualquer brincadeira feita nessas plataformas poderia ser levada a sério.

Usuários das redes sociais também não se mostraram satisfeitos com a medida, chamada, por alguns, de "Big Brother nos Estados Unidos".

O jornalista Peter Clarke, ex-funcionário da rede ABC na Austrália, disse no Twitter que as regras são "ainda mais invasivas" e que outros destinos, com isso, "acenam" para os viajantes.

O também jornalista Kumar Manish, da Índia, recomendou, via Twitter, que os internautas "pensem duas vezes, três vezes antes de postarem qualquer coisa". "Seu sonho americano pode ir por água abaixo com base nisso", acrescentou.

ANEXO D – H4 / Dados

Projeção da população 2018: número de habitantes do país deve parar de crescer em 2047

A população do país deverá crescer até 2047, quando chegará a 233,2 milhões de pessoas. Nos anos seguintes, ela cairá gradualmente, até os 228,3 milhões em 2060. Essas são algumas das informações da revisão 2018 da Projeção de População do IBGE, que estima demograficamente os padrões de crescimento da população do país, por sexo e idade, ano a ano, até 2060.

Em 2060, um quarto da população (25,5%) deverá ter mais de 65 anos. Nesse mesmo ano, o país teria 67,2 indivíduos com menos de 15 e acima dos 65 anos para cada grupo de 100 pessoas em idade de trabalhar (15 a 64 anos).

Em relação à migração internacional, a projeção considerou a emigração da Venezuela para Roraima entre 2015 e 2022. Nesse período, migrariam para o estado cerca de 79,0 mil venezuelanos.

A taxa de fecundidade total para 2018 é de 1,77 filho por mulher. Em 2060, o número médio de filhos por mulher deverá reduzir para 1,66. Entre os estados, as maiores taxas de fecundidade serão em Roraima (1,95), seguido por Pará, Amapá, Maranhão, Mato Grosso, Mato Grosso do Sul, todos com 1,80. As menores deverão ser no Distrito Federal (1,50) e em Goiás, Rio de Janeiro e Minas Gerais, todos com 1,55. Já a idade média em que as mulheres têm filhos é de 27,2 anos em 2018 e deverá chegar a 28,8 anos em 2060.

A revisão 2018 estendeu a Projeção da População para unidades da federação até 2060. Santa Catarina, que hoje tem a maior esperança de vida ao nascer para ambos os sexos (79,7 anos), deverá manter essa liderança até 2060, chegando aos

84,5 anos. No outro extremo, o Maranhão (71,1 anos) tem a menor esperança de vida ao nascer em 2018, condição que deverá ser ocupada pelo Piauí em 2060 (77,0 anos).

A projeção detalha a dinâmica de crescimento da população brasileira, acompanha suas principais variáveis (fecundidade, mortalidade e migrações) e projeta o número de habitantes do Brasil e das 27 unidades da federação, ano a ano, de 2010 a 2060. Esse estudo demográfico é realizado em parceria com órgãos de planejamento de quase todos os estados brasileiros e segue as recomendações da Divisão de População das Nações Unidas. O material de apoio da Projeção da População 2018 está à direita desta página.

Em 2060, população deverá voltar a patamar próximo ao de 2034

A população total projetada para o país em 2018 é de 208,5 milhões. Esse número crescerá até alcançar o máximo de 233,2 milhões em 2047. A partir desse ano, a população irá diminuir até atingir 228,3 milhões em 2060, nível equivalente ao de 2034 (228,4 milhões).

A taxa de fecundidade total (número médio de filhos por mulher) projetada para 2018 é de 1,77 filho por mulher, e deverá reduzir para 1,66 em 2060. A revisão 2018 mostrou que o envelhecimento do padrão da fecundidade é determinado pelo aumento na quantidade de mulheres que engravidam entre 30 e 39 anos e pela redução da participação de mulheres entre 15 e 24 anos na fecundidade em todas as grandes regiões do país.

#pracegover gráfico de barra horizontal demonstrando a projeção de cada estado até 2070

Idade média em que as mulheres têm filhos deverá chegar a 28,8 anos em 2060

Além da queda do nível de fecundidade, projeta-se que o padrão etário de fecundidade por idade da mulher também se altere, conforme já observado nas últimas

décadas, em direção a um envelhecimento da fecundidade no Brasil. A idade média em que as mulheres têm filhos, que está em 27,2 anos em 2018, deverá chegar a 28,8 anos em 2060.

Estima-se que um quarto da população terá mais de 65 anos em 2060

Em 2060, o percentual da população com 65 anos ou mais de idade chegará a 25,5% (58,2 milhões de idosos), enquanto em 2018 essa proporção é de 9,2% (19,2 milhões). Já os jovens (0 a 14 anos) deverão representar 14,7% da população (33,6 milhões) em 2060, frente a 21,9% (44,5 milhões) em 2018.

O envelhecimento afeta a razão de dependência da população, que é representada pela relação entre os segmentos considerados economicamente dependentes (pessoas com menos de 15 e 65 anos ou mais de idade) e o segmento etário potencialmente produtivo (15 a 64 anos), que é a proporção da população que, em tese, deveria ser sustentada pela parcela economicamente produtiva.

A razão de dependência da população em 2018 é de 44%. Esse indicador significa que 44 indivíduos com menos de 15 e com mais de 64 anos dependem de cada grupo de 100 pessoas em idade de trabalhar (15 a 64 anos). Em 2039, a razão de dependência total deverá ser de 51,5%, quando a proporção de jovens (25,7%) e idosos (25,8%) se equivalerá. Essa proporção total deverá aumentar para 67,2% em 2060.

Ainda em 2010, a razão de dependência era de 47,1%, e atingiu seu valor mínimo em 2017 (44,0%). A partir de então, essa proporção voltaria a crescer, chegando, em 2028, a 47,4%, o mesmo nível de 2010.

Roraima deverá ter a maior taxa de fecundidade em 2060: 1,95 filho por mulher

Atualmente, os estados do Norte e Maranhão (1,93), Espírito Santo (1,83), Paraná (1,80) e Goiás (1,79) possuem a taxa de fecundidade total acima da média nacional. Roraima (2,31), Amazonas (2,28), Acre (2,22) e Amapá (2,11) são os únicos que tem a fecundidade acima do nível de reposição em 2018 (2,1 filhos por mulher).

Já em 2060, Roraima deverá continuar com a maior taxa de fecundidade (1,95), seguido por Pará, Amapá, Maranhão, Mato Grosso, Mato Grosso do Sul, todos com 1,80. As menores deverão ser Distrito Federal (1,50) e Goiás, Rio de Janeiro e Minas Gerais, os três com 1,55.

Taxas de fecundidade total estimadas e projetadas 2000/2060

Unidades da Federação	Taxas de fecundidade									
	Estimadas			Projetadas						
	2000	2010	2015	2018	2020	2030	2040	2050	2060	
Brasil	2,32	1,75	1,80	1,77	1,76	1,72	1,69	1,67	1,66	
Norte	3,14	2,21	2,11	2,03	2,00	1,88	1,83	1,81	1,80	
Rondônia	2,55	1,84	1,88	1,88	1,87	1,82	1,79	1,77	1,76	
Acre	3,94	2,45	2,37	2,22	2,14	1,89	1,80	1,77	1,75	
Amazonas	3,43	2,47	2,39	2,28	2,22	2,01	1,89	1,83	1,79	
Roraima	3,50	2,48	2,30	2,31	2,25	2,06	1,99	1,96	1,95	
Pará	3,09	2,14	1,99	1,92	1,89	1,82	1,80	1,80	1,80	

Amapá	3,82	2,55	2,29	2,11	2,04	1,86	1,82	1,80	1,80
Tocantins	2,73	2,03	1,96	1,93	1,91	1,83	1,80	1,79	1,78
Nordeste	2,59	1,82	1,78	1,75	1,73	1,69	1,67	1,66	1,66
Maranhão	3,25	2,12	1,95	1,93	1,90	1,83	1,81	1,80	1,80
Piauí	2,70	1,80	1,77	1,76	1,75	1,71	1,70	1,69	1,69
Ceará	2,60	1,75	1,74	1,70	1,69	1,66	1,65	1,64	1,64
Rio Grande do Norte	2,45	1,73	1,72	1,65	1,65	1,64	1,64	1,64	1,64
Paraíba	2,45	1,76	1,79	1,76	1,74	1,69	1,67	1,65	1,64
Pernambuco	2,46	1,79	1,83	1,76	1,74	1,68	1,66	1,65	1,64
Alagoas	2,90	2,00	1,83	1,76	1,74	1,69	1,68	1,68	1,68
Sergipe	2,68	1,82	1,78	1,74	1,72	1,67	1,66	1,65	1,65
Bahia	2,37	1,73	1,68	1,69	1,68	1,64	1,63	1,62	1,62
Sudeste	2,07	1,63	1,73	1,70	1,70	1,68	1,65	1,63	1,61
Minas Gerais	2,14	1,57	1,62	1,62	1,61	1,60	1,58	1,57	1,55
Espírito Santo	2,17	1,73	1,85	1,83	1,83	1,79	1,75	1,71	1,67
Rio de Janeiro	2,03	1,59	1,75	1,74	1,73	1,68	1,64	1,59	1,55
São Paulo	2,04	1,67	1,77	1,72	1,72	1,70	1,68	1,67	1,65
Sul	2,15	1,63	1,77	1,74	1,74	1,72	1,71	1,69	1,68
Paraná	2,19	1,72	1,80	1,80	1,79	1,76	1,74	1,71	1,68

Santa Catarina	2,10	1,60	1,74	1,74	1,74	1,72	1,71	1,69	1,68
Rio Grande do Sul	2,13	1,56	1,74	1,68	1,68	1,68	1,68	1,68	1,68
Centro-Oeste	2,16	1,77	1,90	1,87	1,85	1,80	1,75	1,69	1,63
Mato Grosso do Sul	2,33	1,92	2,04	2,02	2,01	1,96	1,91	1,85	1,80
Mato Grosso	2,35	1,88	2,08	2,06	2,05	1,99	1,92	1,86	1,80
Goiás	2,02	1,69	1,82	1,79	1,77	1,72	1,66	1,61	1,55
Distrito Federal	2,13	1,65	1,71	1,68	1,67	1,63	1,59	1,54	1,50

Fonte: IBGE. Diretoria de Pesquisas. Coordenação de População e Indicadores Sociais

Cerca de 79,0 mil imigrantes venezuelanos são estimados em Roraima até 2022

Entre 2015 e 2017, 20,5 mil venezuelanos migraram para Roraima, número que deverá aumentar em 185,4%, ao considerar a projeção de 58,5 mil entre 2018 e 2022.

Na elaboração dos parâmetros e hipóteses para a migração internacional para a revisão 2018 da projeção da população, foram monitorados os fluxos migratórios de todas as nacionalidades, chegando-se a conclusão de que as entradas e saídas no território nacional se compensavam com saldo migratório próximo de zero. A única exceção foram os fluxos oriundos da Venezuela, em função do impacto local em Roraima.

Santa Catarina deverá manter a maior esperança de vida ao nascer até 2060

A revisão de 2018 estendeu a projeção da população por unidade da federação até 2060. Santa Catarina, que hoje já é a que tem a maior esperança de vida ao nascer para ambos os sexos (79,7 anos), deverá se manter nessa posição, com 84,5 anos em 2060. No outro extremo, o Piauí deverá ter a menor esperança de vida ao nascer em 2060 (77,0 anos).

Entre os homens, os valores de esperança de vida mais elevados, projetados para 2060, serão observados em Santa Catarina, de 81,5 anos, e Rio Grande do Sul, de 80,9 anos, enquanto os mais baixos serão os do Piauí, de 72,7 anos, e do Pará, de 73,6 anos. Os valores mais altos de esperança de vida feminina também serão em Santa Catarina, de 87,6 anos, seguido pelo Paraná, com 87,0 anos. Rondônia e Roraima experimentarão as mais baixas esperanças de vida entre as mulheres, de 80,3 anos e 80,8 anos, respectivamente.

A esperança de vida ao nascer e taxa de mortalidade infantil para o país têm revisão prevista para 2022, após a realização do Censo Demográfico 2020.

Fonte: Agência de Notícias IBGE (2018)