



**UNIVERSIDADE FEDERAL DO OESTE DO PARÁ
INSTITUTO DE ENGENHARIA E GEOCIÊNCIAS
BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

JONATHAN OLIVEIRA FERNANDEZ

**INTELIGÊNCIA ARTIFICIAL EM DESTAQUE: UM ESTUDO
DA COBERTURA JORNALÍSTICA BRASILEIRA**

**Santarém - Pará
2025**

JONATHAN OLIVEIRA FERNANDEZ

INTELIGÊNCIA ARTIFICIAL EM DESTAQUE: UM ESTUDO DA
COBERTURA JORNALÍSTICA BRASILEIRA

Trabalho de Conclusão de Curso apresentado como requisito para obtenção do grau de Bacharel em Sistemas de Informação, pela Universidade Federal do Oeste do Pará, no Instituto de Engenharia e Geociências.

Orientador: Prof. Dr. Fábio Manoel França Lobato

Santarém - Pará

2025

Dados Internacionais de Catalogação-na-Publicação (CIP)
Sistema Integrado de Bibliotecas – SIBI/UFOPA

- F363i Fernandez, Jonathan Oliveira
 Inteligência artificial em destaque: um estudo da cobertura jornalística brasileira. /
Jonathan Oliveira Fernandez. - Santarém, 2025.
 43 p. : il.
 Inclui bibliografias.
- Orientador: Fábio Manoel França Lobato.
 Trabalho de Conclusão de Curso (Graduação) – Universidade Federal do Oeste do Pará,
Instituto de Engenharia e Geociências, Bacharelado em Sistemas de Informação.
1. Inteligência Artificial. 2. Cobertura Midiática. 3. Regulamentação da IA. 4. Pro-
cessamento de Linguagem Natural. I. Lobato, Fábio Manoel França, *orient.* II. Título.

CDD: 23 ed. 006.3

JONATHAN OLIVEIRA FERNANDEZ

**INTELIGÊNCIA ARTIFICIAL EM DESTAQUE: UM ESTUDO DA
COBERTURA JORNALÍSTICA BRASILEIRA**

Trabalho de Conclusão de Curso apresentado como requisito para obtenção do grau de Bacharel em Sistemas de Informação, pela Universidade Federal do Oeste do Pará, no Instituto de Engenharia e Geociências.

Conceito: 10

Data de aprovação: 21/02/2025

Banca Examinadora:

Prof. Dr. Fábio Manoel França Lobato - Orientador
Universidade Federal do Oeste do Pará

Profa. Dra. Liviane Ponte Rego
Universidade Federal do Oeste do Pará

Prof. Me. Fabrício Almeida do Carmo
Universidade Federal do Oeste do Pará

AGRADECIMENTOS

Primeiramente, agradeço aos meus pais, José e Maria, que me acompanharam desde meus primeiros passos e sempre estiveram comigo em diversos momentos. Toda a dedicação, apoio e compreensão que me forneceram foram essenciais para que eu pudesse chegar até aqui.

Agradeço ao meu orientador, Prof. Dr. Fábio Lobato, que, com sua competência e expertise, me guiou durante minha trajetória acadêmica, com orientações, conselhos e ensinamentos que foram fundamentais para a realização deste trabalho.

Aos meus amigos que fizeram parte do LACA e aos que ainda fazem, agradeço pelo companheirismo e por dividirem comigo momentos de aprendizado, risadas e vários momentos que se transformaram em boas memórias. A companhia de cada um tornou os desafios do dia a dia mais leves.

A todos, meus sinceros agradecimentos por fazerem parte dessa trajetória e por contribuírem, de forma única, para que este trabalho se tornasse realidade.

“Um homem pode ser destruído, mas nunca derrotado.”

Ernest Hemingway

RESUMO

A Inteligência Artificial (IA) tem se integrado cada vez mais em diferentes camadas da sociedade, trazendo benefícios e riscos. Com o crescimento dessa tecnologia, a cobertura midiática sobre o tema também se intensificou. Assim, este trabalho analisa a cobertura midiática da IA no Brasil, identificando os temas e tendências predominantes nas notícias relacionadas à tecnologia. Para isso, foram analisados quatro portais jornalísticos utilizando modelagem de tópicos. Os resultados obtidos revelam que as notícias abordam o impacto da IA no mercado de trabalho, os avanços na saúde, as questões políticas relacionadas à regulamentação e desinformação, além das inovações tecnológicas. Portanto, o estudo fornece uma visão abrangente dos temas destacados pela mídia nos principais portais brasileiros, contribuindo para o estado da arte ao promover insights que podem servir de subsídio para a formulação de políticas públicas e regulamentações relacionadas à tecnologia.

Palavras-chave: Inteligência Artificial, Cobertura Midiática, Regulamentação da IA, Processamento de Linguagem Natural.

ABSTRACT

Artificial Intelligence (AI) has become increasingly integrated into different layers of society, bringing both benefits and risks. With the growth of this technology, media coverage of it has also intensified. This paper analyzes media coverage of AI in Brazil, identifying the predominant themes and trends in news about the technology. To this end, four news portals were analyzed using topic modeling. The results show that the news addresses the impact of AI on the job market, advances in health, political issues involving regulation and disinformation, and technological innovations. Therefore, the study provides a comprehensive view of the topics highlighted by the media on the main Brazilian portals, contributing to the state of the art by promoting insights that can serve as a basis for the formulation of public policies and regulations regarding the technology.

Keywords: Artificial Intelligence, Media Coverage, AI Regulation, Natural Language Processing.

SUMÁRIO

1	INTRODUÇÃO	8
2	TEXTO DO ARTIGO	11
2.1	Introdução	11
2.2	Trabalhos Relacionados	12
2.3	Metodologia	13
2.3.1	Entendimento do negócio	14
2.3.2	Entendimento dos dados	14
2.3.3	Preparação dos dados	15
2.3.4	Modelagem	15
2.3.5	Avaliação e Implantação	15
2.4	Resultados	16
2.4.1	Coleta de dados	16
2.4.2	Modelagem de tópicos	17
2.5	Conclusões	21
3	CONSIDERAÇÕES FINAIS	23
	REFERÊNCIAS	25
	APÊNDICE A – ARTIGO ORIGINAL	28
	APÊNDICE B – TÓPICOS GERAIS	40
	APÊNDICE C – TÓPICOS SEGMENTADOS POR PERÍODOS ANUAIS	41
	ANEXO A – REGISTRO DE PARTICIPAÇÃO NO EVENTO	42

1 INTRODUÇÃO

Este trabalho foi desenvolvido durante a vigência do autor como bolsista de Iniciação Científica no Laboratório de Computação Aplicada. A ideia deste trabalho surgiu com a participação do Prof. Dr. Fábio Manoel França Lobato como consultor do Ministério da Ciência, Tecnologia e Inovação no Grupo de Trabalho em Economia Digital do G20, como especialista em Inteligência Artificial, em conjunto com a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) como parceira de conhecimento.

Durante o período da consultoria, viu-se que o assunto estava ganhando destaque no Brasil, impulsionado pela formulação de políticas públicas para a adoção e regulamentação da tecnologia. Nesse contexto, debates sobre ética, governança de dados e impactos socioeconômicos da IA se tornaram temas centrais nas discussões do país. Assim, surgiu a ideia de analisar como a mídia estava retratando a Inteligência Artificial para a sociedade, buscando compreender as narrativas predominantes, os possíveis vieses e a influência dessas representações na percepção pública sobre a tecnologia.

Além das orientações do Prof. Fábio Lobato, o estudo contou com a participação de coautores de outras universidades, como a Profa. Solange Oliveira Rezende, da Universidade de São Paulo, que atuou na revisão e validação do trabalho com sua *expertise* na área de Inteligência Artificial. Também participou como coautor um aluno membro do Laboratório de Computação Aplicada (LACA), que foi liderado pelo autor do trabalho.

O trabalho ([FERNANDEZ et al., 2024](#)) foi aceito e apresentado pelo autor no dia 21 de novembro de 2024 na 21ª edição do evento Encontro Nacional de Inteligência Artificial e Computacional (ENIAC) realizado em Belém. A apresentação recebeu elogios, sendo bem avaliado no evento. Em virtude desse reconhecimento e alinhamento com as pesquisas do LACA, onde o autor atuou como bolsista de Iniciação Científica, foi decidido que este seria o Trabalho de Conclusão de Curso (TCC).

A Figura 1.1 apresenta uma foto registrada durante o evento, que ilustra o momento de apresentação do trabalho.

Como apresentado na Figura 1.1, a participação no evento foi uma experiência enriquecedora pela oportunidade de apresentar o trabalho e interagir com especialistas e profissionais da área. O evento proporcionou uma troca de conhecimento a respeito de novas perspectivas sobre o impacto da Inteligência Artificial na sociedade.

A recepção do público em relação ao trabalho foi extremamente positiva, resultando



Figura 1.1 – Foto da apresentação do artigo

em *feedbacks* positivos da plateia, reforçando a relevância do tema abordado. Após a apresentação, algumas perguntas foram feitas pelo *chair* da sessão e por participantes, como:

- “*Vocês procuraram blogs de tecnologia ou outras fontes além dos jornais?*”
 - **Resposta:** Nosso estudo focou em portais jornalísticos conhecidos pelo público brasileiro, portanto, blogs especializados em tecnologia não foram incluídos, mas podem ser explorados em trabalhos futuros.
- “*Tentaram realizar uma análise de sentimento dessas notícias?*”
 - **Resposta:** Uma das ideias que tivemos foi essa, porém, devido ao prazo de submissão do artigo não foi possível realizar essa análise, mas podemos abordar futuramente.
- “*Os tópicos foram anotados com alguma ferramenta?*”
 - **Resposta:** Os tópicos foram anotados manualmente verificando o conjunto de palavras gerado pelo BERTopic.

Essas questões fomentaram um debate produtivo e contribuíram para enriquecer ainda mais a discussão sobre o impacto da tecnologia na sociedade, além de contribuir para trabalhos futuros.

O trabalho está organizado como segue: a Seção 2 apresenta o texto original do artigo, enquanto a Seção 3 é dedicada às considerações finais.

2 TEXTO DO ARTIGO

As subseções a seguir apresentam fielmente o artigo, mantendo a estrutura e os conceitos originais, com o objetivo de proporcionar uma visão detalhada dos métodos, resultados e conclusões apresentados.

2.1 INTRODUÇÃO

O avanço tecnológico tem contribuído para simplificar cada vez mais as tarefas humanas, reduzindo a necessidade de esforço manual em várias atividades e, em alguns casos, até eliminando a necessidade de realizá-las, graças ao uso de algoritmos de Inteligência Artificial (IA) por empresas, governos e organizações (NTOUTSI et al., 2020). O uso dessa tecnologia oferece diversas vantagens que favorecem a tomada de decisões, uma vez que a capacidade de analisar grandes volumes de dados de forma rápida pode ser benéfica para a otimização de processos, a identificação de tendências e o aprimoramento da eficiência operacional (STROPARO et al., 2024).

Um dos avanços mais relevantes na área nos últimos anos tem sido o desenvolvimento e uso dos *Large Language Models* (LLMs), amplamente conhecidos como IA generativa, nos quais os algoritmos são capazes de produzir conteúdo textual, visual e auditivo de maneira quase autônoma (LONGONI et al., 2022). Essa digitalização dos meios de comunicação impactou questões que antes eram consideradas exclusivamente humanas, uma vez que os algoritmos estão auxiliando os humanos em tarefas como a redação de textos, e-mails e livros (PEÑA-FERNÁNDEZ et al., 2023). No entanto, apesar do grande potencial da IA generativa, sua implementação ainda enfrenta uma série de desafios, como a discriminação racial na geração de imagens¹, em que os modelos de IA podem reproduzir e amplificar preconceitos existentes se forem treinados com dados que contêm viés (GARCIA, 2020). Além disso, a técnica *deepfake*, que, embora traga benefícios, pode ser utilizada para gerar desinformação tanto no âmbito social quanto político (MOLINA; BERENGUEL, 2022).

Desse modo, um objetivo crítico para o desenvolvimento da IA é a ampliação da confiança pública, uma vez que a população desempenha um papel fundamental na formulação de políticas públicas e na supervisão regulatória dessa tecnologia. A opinião popular pode influenciar tanto o tipo de IA que será desenvolvida quanto a forma como ela será regulamentada pelo governo (OUCHCHY et al., 2020). Assim, à medida que a IA generativa avança, um número crescente de notícias surge, explorando como os

¹Disponível em: <<http://glo.bo/3M9fsat>>

efeitos dessa evolução impactam nossas vidas (CHOI, 2024). Nesse sentido, a mídia jornalística desempenha um papel preponderante ao traduzir conhecimentos complexos em informações acessíveis, e durante esse processo, influencia a percepção pública sobre a tecnologia, especialmente quando as pessoas se sentem desconfiadas ou ansiosas (CHUAN, 2023).

Considerando que as percepções do público sobre o tema da IA são moldadas pela maneira como a mídia apresenta essas informações, o objetivo deste trabalho é analisar a cobertura midiática a respeito desse tema e entender como os principais portais jornalísticos brasileiros relatam as notícias, quais temas são predominantes e de que maneira essa narrativa pode impactar a opinião pública sobre a tecnologia. Para essa finalidade, foram utilizadas técnicas de Processamento de Linguagem Natural (PLN) e modelagem de tópicos com a técnica BERTopic, visando identificar padrões temáticos recorrentes nas notícias e compreender como a mídia constrói e dissemina essas narrativas. Ao realizar essa análise, é possível observar quais assuntos recebem maior destaque e os possíveis impactos éticos, econômicos e tecnológicos, contribuindo assim para o debate sobre a regulação dessa tecnologia emergente.

2.2 TRABALHOS RELACIONADOS

No trabalho (OUCHCHY et al., 2020), os autores analisaram a cobertura midiática da IA com enfoque nas questões éticas relacionadas à tecnologia. Utilizando o banco de dados *NexisUni*, os pesquisadores coletaram os dados no período de 2014 a 2018, resultando em uma amostra de 254 publicações, incluindo notícias, artigos de revistas e blogs. As publicações foram categorizadas em cinco áreas: questões éticas; princípios baseados em estruturas éticas; recomendações das publicações; tom da publicação; tipo de tecnologia específica. Os resultados revelaram que inicialmente as reportagens possuíam um tom otimista e entusiasmado, que se tornou mais crítico ou neutro ao longo dos anos. Os temas éticos mais discutidos foram preconceito e proteção de dados, enquanto as recomendações ainda estão em estágios iniciais. O trabalho sugere que a mídia mantém uma abordagem prática na cobertura da ética da IA, mas muitos autores carecem de conhecimento adequado sobre o tema. Em contraste com o presente estudo, o trabalho correlato não menciona a utilização de métodos automatizados para análise dos dados (*e.g.* modelagem de tópicos), sendo este um diferencial deste trabalho.

O estudo de (COOLS et al., 2024) examinou o panorama midiático sobre IA e automação em jornais dos Estados Unidos, analisando a forma que a mídia apresenta os benefícios e riscos associados a estes temas. Por meio da *Application Programming Interface* (API) dos portais *New York Times* (NYT) e *The Washington Post* (WP), os pesquisadores coletaram notícias de 1985 a 2020. Posteriormente, realizaram modelagem de tópicos com

Latent Dirichlet Allocation (LDA) e análise manual de enquadramento para a análise dos dados, na qual rotularam as notícias como neutras, positivas e negativas em uma escala de utópico a distópico. Os resultados revelaram dez tópicos e quadros distintos. Os tópicos abordaram temas como Política, Educação, e Esporte, com oito tópicos possuindo relevância até um determinado período, e dois que permanecem relevantes ao longo do tempo, nos quais observou-se um aumento a partir de 2013 nos tópicos Trabalho, Arte e Educação. Em relação aos enquadramentos, a mídia tem sido mais otimista do que pessimista ao longo do tempo, embora ainda existam narrativas distópicas que retratam a IA e a automação como tecnologias que irão substituir os humanos no futuro. Portanto, o estudo apresenta uma visão geral de longo prazo sobre estas tecnologias, revelando tanto temas positivos, quanto negativos. Apesar do estudo correlato também utilizar modelagem de tópicos, o presente trabalho se diferencia ao empregar a técnica BERTopic, que permite uma análise mais detalhada ao incorporar informações contextuais na formação dos tópicos, cuja eficácia foi comprovada por (EGGER; YU, 2022).

O trabalho de (SUN et al., 2020) realizou uma análise na cobertura midiática sobre IA ao longo das últimas quatro décadas, utilizando como fontes os veículos de notícias NYT, WP, *The Guardian* e *USA Today*. Os autores analisaram 1.776 notícias usando *K-Means* para identificar tópicos e temas principais, análise de léxico para detectar padrões de argumentação, análise de rede para mapear países e organizações mencionadas e análise de correspondência múltipla para explorar associações entre os dados. Os resultados revelaram 14 *clusters*, dos quais os pesquisadores extraíram as palavras-chave mais relevantes de cada tema, revelando tópicos como “Assistência médica”, “Emprego e economia”, “Regulamentação e política”, entre outros. Na análise de léxico identificaram nove padrões de argumentação, no qual o mais prevalente foi o pragmático, relacionado a notícias sobre aplicação prática da IA no cotidiano. A análise de rede destacou potências como Estados Unidos, China, Rússia, Japão e Alemanha, além das organizações mais citadas como Google, Facebook e Microsoft. Por fim, a análise de correspondência múltipla revelou que o NYT tende a focar mais em notícias sobre negócios, enquanto o *The Guardian* se concentra em assuntos sobre robôs e algoritmos; os demais portais estão relacionados às notícias sobre aplicações da IA. Assim, o estudo conclui que embora a mídia aborde IA há décadas, a cobertura se intensificou somente nos últimos cinco anos, adotando uma abordagem mista e equilibrada.

2.3 METODOLOGIA

Seguindo o guia proposto por (WIRTH; HIPPEL, 2000), este estudo utilizou a metodologia *Cross-Industry Standard Process for Data Mining* (CRISP-DM), amplamente reconhecida na literatura para conduzir análises de dados e mineração de textos, conforme

demonstrado em estudos como (SCHNEIDER et al., 2023; COSTA et al., 2022). O CRISP-DM é composto por seis etapas: Entendimento do negócio, Entendimento dos dados, Preparação dos dados, Modelagem, Avaliação e Implantação; detalhadas nas subseções a seguir.

2.3.1 Entendimento do negócio

A análise de notícias sobre IA foi enriquecida pela experiência das pessoas autoras e reuniões com especialistas da área, que incluíram debates com membros do Grupo de Trabalho em Economia Digital do G20, participações em eventos como o *AI Summit* e colaborações com acadêmicos e profissionais de diferentes setores. A combinação da *expertise* das pessoas autoras e do conhecimento obtido com outros especialistas possibilitou uma compreensão mais profunda sobre as implicações sociais, econômicas e éticas da IA, proporcionando a base necessária para este trabalho.

Paralelamente foi realizado um levantamento da literatura do tipo *Ad Hoc* (RALPH; BALTES, 2022), visando compreender a importância da análise de notícias, uma vez que a forma como essa tecnologia é retratada pode afetar sua aceitação e a formação de políticas públicas, no qual os principais estudos encontrados foram apresentados tanto na Seção 2.1 quanto na Seção 2.2.

2.3.2 Entendimento dos dados

Esta etapa foi dividida em duas fases: seleção de palavras-chave e coleta de dados. Para a primeira fase foi utilizado o método *Delphi* apresentado no estudo de (BAINES; BERE, 2018), que consiste na concordância entre especialistas, onde as palavras-chave são listadas e validadas, resultando nas seguintes palavras-chave: “Inteligência Artificial”, “OpenAI”, “Google Gemini” e “Microsoft Copilot”.

Na fase de coleta de dados os portais jornalísticos selecionados foram o Portal G1, Metrôpoles, Folha de São Paulo, e CNN Brasil devido à popularidade e abrangência na cobertura de notícias destes portais, bem como a diversidade editorial e a credibilidade consolidada no jornalismo brasileiro. Os dados foram coletados por meio de *web scraping* utilizando um *web crawler* e, posteriormente, salvos em formato *Comma-Separated Values* (CSV) para facilitar a manipulação e análise eficiente. Os atributos coletados incluem: 1) título da notícia; 2) resumo; 3) corpo do texto; 4) data da publicação; 5) região; 6) link da notícia; 7) palavra-chave; e 8) plataforma de notícia.

2.3.3 Preparação dos dados

Esta etapa também foi dividida em duas fases: filtragem e pré-processamento. Utilizando como base a teoria fundamentada descrita por (BAUMER et al., 2017), foi realizada uma filtragem para identificar quais notícias se enquadravam no escopo do estudo, já que muitas delas apenas mencionavam IA de passagem, sem discutir o tema de maneira substancial. O processo envolveu a leitura dos títulos e conteúdos das notícias, focando na relação da notícia com a palavra-chave, a fim de garantir a consistência dos dados e mitigar possíveis inconsistências. As notícias que não abordavam diretamente o assunto foram removidas do conjunto de dados.

Em seguida, os textos passaram para a fase de pré-processamento, visando a remoção de ruídos, palavras irrelevantes, pontuações e acentos que podem afetar a qualidade e relevância das informações extraídas. Para isso foram aplicados *scripts* escritos em *Python* seguindo o *pipeline* de (ARAÚJO et al., 2024), que envolve os seguintes passos: i) padronização para caixa-baixa; ii) remoção de pontuação, numerais e caracteres especiais; iii) remoção de *stopwords*¹.

2.3.4 Modelagem

Nesta etapa, foram conduzidas análises com o objetivo de identificar temas recorrentes nas notícias, utilizando técnicas de PLN e modelagem de tópicos com o BERTopic. Essa técnica cria tópicos a partir de *embeddings* gerados por modelos de linguagem pré-treinados. O BERTopic gera *embeddings* contextuais para os textos, capturando o significado semântico das palavras em diferentes contextos. Em seguida, realiza a clusterização dos textos que possuem semântica semelhante, e os tópicos são extraídos utilizando a medida *Class-Based Term Frequency-Inverse Document Frequency* (c-TF-IDF), que considera a frequência e o contexto dos termos no corpus do documento (GROOTENDORST, 2022).

Para a análise, foi empregada a configuração padrão do BERTopic, utilizando o modelo *Multilingual*, que possui suporte para a língua portuguesa. A fim de proporcionar uma visão geral e detalhada, optou-se por segmentar os tópicos por portal jornalístico, selecionando os temas mais distintos entre eles. Posteriormente, os tópicos gerados foram anotados manualmente para uma melhor interpretação dos resultados.

2.3.5 Avaliação e Implantação

A etapa de Avaliação manteve a abordagem da teoria fundamentada, no qual uma amostra representativa dos dados foi separada com base em um intervalo de confiança de 95% e uma taxa de erro de 10%, para que os autores realizassem a leitura dessa amostra e em

¹Palavras que não têm relevância, mas aparecem com frequência (*e.g.* preposições, artigos, pronomes).

seguida confrontassem com os resultados obtidos pelos algoritmos, objetivando a detecção de incongruências ou informações incorretas que pudessem ter passado despercebidas nas etapas anteriores. Posteriormente esses resultados foram apresentados ao grupo de pesquisa e avaliados em conjunto utilizando a técnica *ground truth*, além da discussão com os especialistas do domínio que participaram do entendimento do negócio.

A etapa de implantação é apoiada por um relatório técnico contendo os resultados obtidos ao longo do estudo, visando proporcionar *insights* relevantes para subsidiar o planejamento de políticas públicas.

2.4 RESULTADOS

Nesta seção são apresentados e discutidos os resultados obtidos considerando a metodologia descrita na seção anterior. Primeiramente são apresentados os resultados referentes à coleta de dados, seguidos pelos resultados obtidos com a modelagem de tópicos.

2.4.1 Coleta de dados

Foram coletadas 12.308 notícias publicadas nos portais jornalísticos no período de 2016 a 2024. Após a aplicação do pré-processamento e da filtragem, restaram 2.557 notícias, as quais foram utilizadas para análises posteriores. Por meio da análise exploratória do conjunto de dados foi possível obter a distribuição da quantidade de notícias por plataforma, conforme apresentado na Figura 2.1.

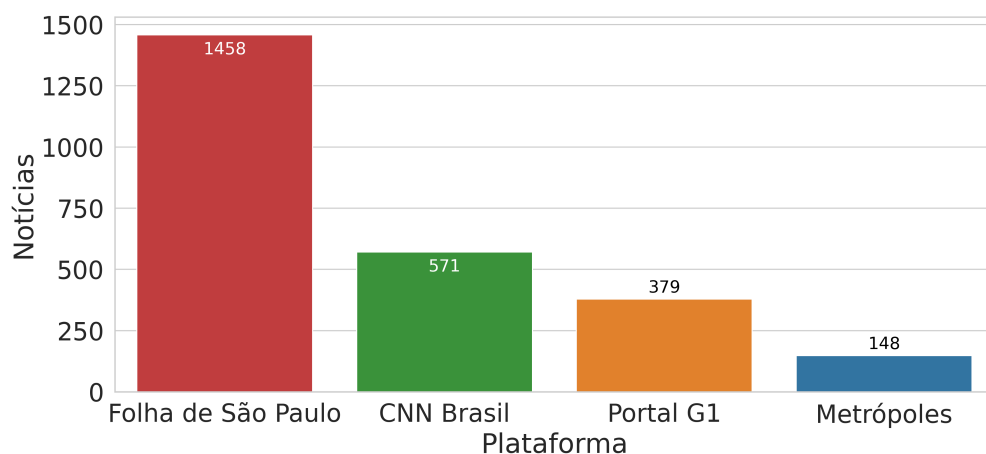


Figura 2.1 – Quantidade de notícias por plataforma.

Vide a Figura 2.1 é possível observar que o portal Folha de São Paulo possui o maior número de notícias, o que pode ser atribuído ao fator de que o jornal é um dos maiores veículos de comunicação no Brasil, onde possui características como textos de qualidade e conteúdos diferentes dos tabloides, no qual priorizam tanto a cobertura da mídia nacional quanto a local (*i.e.* São Paulo) (MARQUES et al., 2020). Além dos resultados mencionados,

o volume de publicações anuais nas plataformas também apresentou variações ao longo dos anos analisados, refletindo mudanças no cenário jornalístico sobre IA. A Figura 2.2 apresenta a distribuição anual de notícias publicadas durante o período analisado.

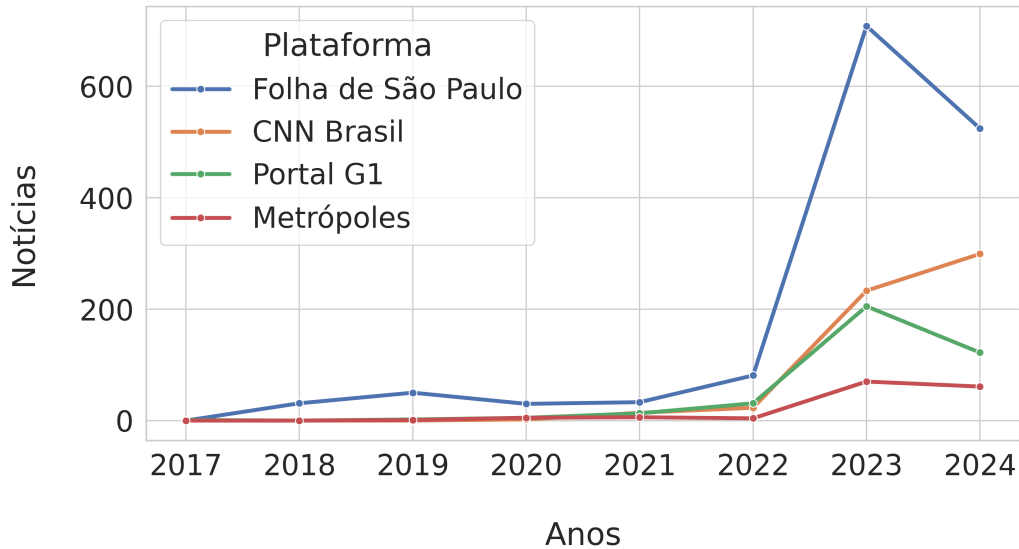


Figura 2.2 – Quantidade de notícias por ano.

De acordo com a Figura 2.2, as notícias sobre esse tema começaram a ter um aumento a partir do ano de 2022, indicando um crescente interesse da mídia sobre o assunto. Essa intensificação pode ser explicada pelo fato de que 2022 foi o ano em que o ChatGPT foi lançado e começou a ganhar visibilidade, causando um ponto de inflexão no campo da IA, gerando discussões sobre suas aplicações, implicações e impactos. Adicionalmente, a Figura 2.3 apresenta uma nuvem de palavras que destaca os termos mais frequentes encontrados nas notícias.

Por meio da análise da Figura 2.3 é possível identificar diversas informações relevantes, como as grandes corporações que atuam na área, além de *startups* e outras empresas de tecnologia. A presença de termos como “foto”, “video” e “texto” demonstram a relevância do processamento e geração de mídia, o que pode se referir a tecnologias de *deepfake* e outras formas de manipulação de conteúdo, onde o termo “caso” pode estar ligado a estes incidentes de manipulação. O destaque em “trabalho” e “pesquisa” indicam a relevância da IA no mercado de trabalho e nos esforços acadêmicos para impulsionar o avanço da tecnologia. Por outro lado, termos como “segurança”, “risco” e “governo” sugerem uma preocupação com os aspectos éticos e regulatórios da IA, evidenciando a necessidade de políticas para o uso responsável da IA.

2.4.2 Modelagem de tópicos

Os resultados da modelagem de tópicos revelaram 8 tópicos individuais. A Tabela 2.1 apresenta os principais assuntos abordados nas publicações, segmentados por

os setores, com alguns experimentando transformações mais profundas que outros. Por exemplo, o setor de finanças foi altamente impactado pela automação de tarefas como detecção de fraudes e análise de riscos.

O setor de saúde tem apresentado avanços promissores na melhoria de diagnósticos e atendimentos ao paciente, evidenciando o potencial transformador da IA em diferentes áreas. O que está ligado ao tópico “Futuro da saúde” que aborda notícias sobre os avanços da IA na saúde e bem-estar da sociedade, abrangendo as inovações tecnológicas que impulsionam este setor, como o monitoramento de pacientes gestantes¹. No entanto, (BUNZ; BRAGHIERI, 2022) informam que as manchetes devem ser cuidadosas ao noticiar as contribuições da IA para a saúde, por haver o risco de descrever os sistemas de IA como algo superior ao humano, antropomorfizando a tecnologia. Esse cuidado é reforçado por estudos como o de (SPATOLA; URBANSKA, 2020), que identifica uma tendência crescente em atribuir expectativas positivas, e até mesmo divinas na IA, o que pode afetar a percepção do público sobre o papel dessa tecnologia.

O portal CNN Brasil apresenta destaque na cobertura sobre a área mercadológica da IA, apresentando tópicos sobre o atual mercado e a introdução de novos recursos inteligentes nos dispositivos. No tópico “Mercado tecnológico” é possível observar por meio dos termos a cobertura sobre as grandes empresas como OpenAI e Microsoft, explorando suas inovações e estratégias de mercado, visto que estas empresas são parceiras de mercado². As notícias reportam como as empresas estão integrando recursos inteligentes em seus produtos e serviços³, um exemplo notável destas inovações é o uso da tecnologia do ChatGPT no mecanismo de busca do navegador Edge, no qual a IA é utilizada para oferecer melhores resultados de busca (SIEGLE, 2023).

O tópico “Dispositivos inteligentes” também indica uma cobertura no lançamento de dispositivos e tecnologias baseadas em IA, como a *Apple Intelligence*, um novo modelo de IA que está sendo desenvolvido para aprimorar a experiência do usuário em dispositivos como *iPhone*, *iPad* e *Mac*, incluindo funcionalidades para escrita, geração de imagens e integrações com a assistente virtual Siri (MOHAMED; LUCKE-WOLD, 2024). Além disso, a cobertura inclui inovações do Google, como os chips *Tensor Processing Unit* (TPU) projetados para otimizar o desempenho da IA nas tarefas de aprendizado de máquina, PLN e entre outros. Notavelmente, a Apple tem integrado esses chips TPUs em seus processos de treinamento de modelos de IA⁴, refletindo o impacto dos desenvolvimentos inovadores

¹Disponível em: <<https://bit.ly/3Ai3G7>>

²Disponível em: <<https://bit.ly/microsoft-partnerships>>

³Disponível em: <<https://bit.ly/microsoft-openai>>

⁴Disponível em: <<https://bit.ly/apple-chips-google>>

do Google no campo de IA.

A cobertura do portal G1 está inclinada para um aspecto político, abordando temas como eleições e a regulamentação da IA na Europa. O tópico “Política e IA”, foca em como a IA está influenciando os processos eleitorais, desde o uso para as campanhas até a disseminação de desinformação por meio de *deepfakes* nas redes sociais¹, uma vez que os conteúdos manipulados podem ser utilizados para enganar eleitores e influenciar a opinião pública. De acordo com (MOLINA; BERENGUEL, 2022), o uso da tecnologia para fins políticos gera preocupações, podendo ser uma ameaça à democracia. Assim, o Brasil tem avançado na formulação de políticas para lidar com os impactos da IA, com a criação do Projeto de Lei (PL) que visa regulamentar o desenvolvimento e o uso da IA (PL 2.338/2023), refletindo um esforço para garantir que a tecnologia seja usada de forma ética e segura.

Assim como na União Europeia (UE), que foi a região pioneira na regulamentação da IA. O tópico “Regulamentação da IA” está associado as notícias que tratam sobre como a UE visa moldar a regulamentação global de tecnologia, replicando o sucesso do Regulamento Geral de Proteção de Dados (GDPR), considerado um modelo de alta qualidade para as leis de proteção de dados, no qual vários países basearam suas estruturas de leis nesse modelo (BELLI et al., 2023).

No portal Metrópolis os tópicos estão associados à manipulação de dados e assistentes virtuais, como *chatbots*. O tópico “Manipulação digital” está relacionado às notícias sobre como os dados são processados e utilizados para a disseminação de desinformação, desde notícias falsas até *deepfakes*, onde muitas dessas práticas envolvem celebridades, políticos² e diretores executivos de empresas, uma vez que o fácil acesso de fotos e vídeos na internet facilitam a criação e circulação desses conteúdos (MOLINA; BERENGUEL, 2022). Além disso, estas práticas também podem ser utilizadas para causar danos a outras pessoas, ocasionando em consequências graves como a difamação e invasão de privacidade.

O tópico “Manipulação Digital” do portal Metrópolis apresenta similaridade com o tópico “Política e IA” abordado pelo portal G1. Ambos os tópicos destacam como a manipulação de dados e a propagação de desinformação são temas preocupantes no cenário atual, ilustrando as consequências dessas práticas e ressaltando a necessidade de impôr regras éticas para o uso e desenvolvimento dessas tecnologias, visando mitigar os impactos negativos na sociedade.

¹Disponível em: <<http://glo.bo/3Mnm4Cj>>

²Disponível em: <<https://bit.ly/deepfake-politicos>>

Por fim, o tópico “Chats inteligentes” explora o desenvolvimento e aplicação de *chatbots* e outras tecnologias de IA. (MAEDCHE et al., 2019) apontam que os assistentes virtuais estão se tornando cada vez mais comuns e integrados ao cotidiano dos usuários, por exemplo, assistentes que utilizam comando de voz como a Amazon Alexa ou interfaces baseadas em texto para o atendimento ao cliente. Ainda, os autores apresentam preocupações em torno dessas tecnologias, gerando debates sobre fraudes e enganos, visto que elas apresentam semelhanças com a linguagem humana.

Os tópicos apresentados e segmentados por portais de notícias revelam a amplitude da cobertura jornalística sobre a IA, apresentando as transformações em várias áreas, os desafios políticos e questões sobre a ética e regulação da tecnologia. Essa análise demonstra o acompanhamento da mídia sobre a evolução da IA, refletindo o papel central da imprensa em conscientizar o público sobre os benefícios e riscos associados a tecnologia, contribuindo para a percepção do público sobre a IA.

2.5 CONCLUSÕES

Os avanços da IA têm trazido inúmeras facilidades e benefícios para a sociedade, mas também apresentam desafios e riscos significativos onde a falta de ética, supervisão regulatória adequada e responsabilidade podem resultar em consequências maiores. Nesse sentido, a mídia jornalística atua como uma ferramenta de divulgação e discussão sobre esses desafios, promovendo uma compreensão crítica da população a respeito do progresso de tecnologias emergentes como a IA.

Nesse contexto, o presente trabalho teve como objetivo analisar a cobertura midiática da IA no Brasil, investigando como os diferentes portais de notícias abordam a IA e suas implicações. Foram analisadas 2.557 notícias por meio de uma modelagem de tópicos, no qual os resultados revelaram temas como mercado de trabalho, saúde, política, manipulação de conteúdo e inovações na área. O estudo fornece uma visão geral de quais temas estão sendo destacados na mídia, oferecendo contribuições para o estado da arte ao apresentar como a IA é retratada; fornecendo subsídios para debates sobre os desafios éticos, sociais e econômicos advindos dessa tecnologia.

Para trabalhos futuros, pretende-se aumentar a quantidade de análises no conjunto de dados, como análise de sentimentos ou rede semânticas, para oferecer uma perspectiva mais detalhada. Ademais, a utilização de LLMs como ferramenta de análise pode permitir uma interpretação mais rica das narrativas sobre a IA na mídia jornalística, identificando nuances linguísticas e contextuais. Outra sugestão de trabalho futuro é explorar a etapa de implantação do CRISP-DM, com um relatório técnico contendo os resultados obtidos ao longo do estudo, a fim de proporcionar *insights* relevantes para subsídio de políticas públicas,

como a regulação da IA, de forma que recurso seja conduzido de maneira responsável, visando o bem-estar da sociedade como um todo.

AGRADECIMENTOS

Este trabalho foi apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) DT-303031/2023-9, PQ-309575/2021-4, PIBIC - 166028/2024-0; e pela Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) PRONEM-FAPESPA/CNPq nº 045/2021.

3 CONSIDERAÇÕES FINAIS

A relevância deste trabalho para a área de Sistemas de Informação está na conexão entre a tecnologia, sociedade e tomada de decisão baseada em dados. Visto que à medida em que a Inteligência Artificial se torna um tema central em debates políticos, econômicos e sociais, entender como a mídia retrata essa tecnologia se torna importante para avaliar sua aceitação, desafios e impactos na sociedade.

A escolha de utilizar um artigo científico como TCC é devido ao formato ser amplamente reconhecido na academia, o que contribui para a visibilidade do estudo, além de possuir validação tanto dos autores do trabalho quanto do evento. Essa experiência foi importante para o meu crescimento acadêmico, pois fortaleceu meu currículo e teve um impacto direto no meu ingresso no mestrado em Engenharia de Computação e Sistemas na Universidade Estadual do Maranhão (UEMA). O trabalho também me permitiu desenvolver habilidades, como pesquisa, análise crítica e escrita científica, que foram decisivas para minha aprovação.

Dessa maneira, este estudo se destaca ao demonstrar a importância das técnicas de Processamento de Linguagem Natural na extração de conhecimento a partir de grandes volumes de dados textuais. O uso de modelagem de tópicos permitiu identificar padrões na comunicação sobre IA, contribuindo para um entendimento mais estruturado de como essa tecnologia é discutida nos meios de comunicação. A abordagem adotada reforça o papel da governança de dados e da ética no desenvolvimento e aplicação de sistemas inteligentes, uma vez que a forma como a IA é percebida pode influenciar diretamente sua aceitação e o desenvolvimento de regulamentações mais justas e eficientes.

Outro ponto relevante deste estudo é seu impacto como subsídio para a criação de estratégias mais eficazes de comunicação e políticas públicas alinhadas com a realidade tecnológica, possibilitando que formuladores de políticas, empresas e pesquisadores ajustem suas abordagens para mitigar medos infundados, esclarecer conceitos e promover um debate mais informado e equilibrado.

Como trabalhos futuros, está planejada a expansão do *web crawler* para incluir portais jornalísticos internacionais, como BBC, The New York Times e El País, que permitirá uma análise comparativa entre diferentes regiões e contextos culturais, possibilitando a identificação de padrões globais e variações na cobertura midiática sobre Inteligência Artificial. Com isso, será possível avaliar como fatores econômicos, políticos e culturais influenciam a construção da narrativa sobre essa tecnologia, enriquecendo ainda mais a

compreensão sobre sua recepção no cenário global. Ainda, pretende-se expandir as análises no conjunto de dados, implementando análise de sentimentos e redes semânticas.

Este estudo abre caminho para pesquisas futuras que explorem a interseção entre Inteligência Artificial, comunicação e percepção pública, incentivando abordagens interdisciplinares que conectem tecnologia, linguística, ciência de dados e ciências sociais. Assim, é possível compreender melhor o impacto da IA na sociedade, e contribuir para um debate mais transparente e embasado sobre o papel dessa tecnologia no mundo contemporâneo.

REFERÊNCIAS

- ARAÚJO, G. D. S.; FERNANDEZ, J. O.; SILVA, M. S. D.; LOBATO, F. M. F. Persona and issue analysis on tourism social media: a case study of pirenópolis, goiás, brazil. In: **Proceedings of the 20th Brazilian Symposium on Information Systems**. [S.l.: s.n.], 2024. p. 1–10.
- BAINES, R. L.; BERE, S. Regan de. Optimizing patient and public involvement (ppi): Identifying its “essential” and “desirable” principles using a systematic review and modified delphi methodology. **Health Expectations**, Wiley Online Library, v. 21, n. 1, p. 327–335, 2018.
- BAUMER, E. P.; MIMNO, D.; GUHA, S.; QUAN, E.; GAY, G. K. Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? **Journal of the Association for Information Science and Technology**, Wiley Online Library, v. 68, n. 6, p. 1397–1410, 2017.
- BELLI, L.; CURZI, Y.; GASPAR, W. B. Ai regulation in brazil: Advancements, flows, and need to learn from the data protection experience. **Computer Law & Security Review**, Elsevier, v. 48, p. 105767, 2023.
- BUNZ, M.; BRAGHIERI, M. The ai doctor will see you now: assessing the framing of ai in news coverage. **AI & society**, Springer, v. 37, n. 1, p. 9–22, 2022.
- CHOI, S. Temporal framing in balanced news coverage of artificial intelligence and public attitudes. **Mass Communication and Society**, Taylor & Francis, v. 27, n. 2, p. 384–405, 2024.
- CHUAN, C.-H. A critical review of news framing of artificial intelligence. **Handbook of Critical Studies of Artificial Intelligence**, Edward Elgar Publishing, p. 266–276, 2023.
- COOLS, H.; GORP, B. V.; OPGENHAFFEN, M. Where exactly between utopia and dystopia? a framing analysis of ai and automation in us newspapers. **Journalism**, SAGE Publications Sage UK: London, England, 2024.
- COSTA, G. d. S.; COUTO, D. C.; JUNIOR, A. F. J.; LOBATO, F. M. Feminismo e redes sociais online: uma análise de tweets sobre o dia internacional da mulher. In: SBC. **Anais do XI Brazilian Workshop on Social Network Analysis and Mining**. [S.l.], 2022. p. 169–180.
- EGGER, R.; YU, J. A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. **Frontiers in sociology**, Frontiers Media SA, v. 7, p. 886498, 2022.
- FERNANDEZ, J.; PONTES, A.; REZENDE, S.; LOBATO, F. Artificial intelligence in the spotlight: A study of brazilian news coverage. In: **Anais do XXI Encontro Nacional de Inteligência Artificial e Computacional**. Porto Alegre, RS, Brasil: SBC, 2024. ISSN 2763-9061.

- GARCIA, A. C. Ética e inteligencia artificial. **Computação Brasil**, n. 43, p. 14–22, 2020.
- GEORGE, A. S.; GEORGE, A. H.; MARTIN, A. G. Chatgpt and the future of work: a comprehensive analysis of ai's impact on jobs and employment. **Partners Universal International Innovation Journal**, v. 1, n. 3, p. 154–186, 2023.
- GROOTENDORST, M. Bertopic: Neural topic modeling with a class-based tf-idf procedure. **arXiv preprint arXiv:2203.05794**, 2022.
- LONGONI, C.; FRADKIN, A.; CIAN, L.; PENNYCOOK, G. News from generative artificial intelligence is believed less. In: **Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency**. [S.l.: s.n.], 2022. p. 97–106.
- MAEDCHE, A.; LEGNER, C.; BENLIAN, A.; BERGER, B.; GIMPEL, H.; HESS, T.; HINZ, O.; MORANA, S.; SÖLLNER, M. Ai-based digital assistants: Opportunities, threats, and research perspectives. **Business & Information Systems Engineering**, Springer, v. 61, p. 535–544, 2019.
- MARQUES, F. P. J.; MIOLA, E.; MITOZO, I.; MONT'ALVERNE, C. Similar, but not the same: Comparing editorial and news agendas in brazilian newspapers. **Journalism Practice**, Taylor & Francis, v. 14, n. 9, p. 1066–1086, 2020.
- MOHAMED, A. A.; LUCKE-WOLD, B. Apple intelligence in neurosurgery. **Neurosurgical Review**, Springer, v. 47, n. 1, p. 327, 2024.
- MOLINA, A. C.; BERENGUEL, O. L. Deepfake: A evolução das fake news. **Research, Society and Development**, v. 11, n. 6, p. e56211629533–e56211629533, 2022.
- NTOUTSI, E.; FAFALIOS, P.; GADIRAJU, U.; IOSIFIDIS, V.; NEJDL, W.; VIDAL, M.-E.; RUGGIERI, S.; TURINI, F.; PAPADOPOULOS, S.; KRASANAKIS, E. et al. Bias in data-driven artificial intelligence systems—an introductory survey. **Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery**, Wiley Online Library, v. 10, n. 3, p. e1356, 2020.
- OUCHCHY, L.; COIN, A.; DUBLJEVIĆ, V. Ai in the headlines: the portrayal of the ethical issues of artificial intelligence in the media. **AI & SOCIETY**, Springer, v. 35, p. 927–936, 2020.
- PEÑA-FERNÁNDEZ, S.; MESO-AYERDI, K.; LARRONDO-URETA, A.; DÍAZ-NOCI, J. Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media. **El Profesional de la información**, v. 32, n. 2, 2023.
- RALPH, P.; BALTES, S. Paving the way for mature secondary research: the seven types of literature review. In: **Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering**. [S.l.: s.n.], 2022. p. 1632–1636.
- SCHNEIDER, J.; SEIDEL, S.; BASALLA, M.; BROCKE, J. vom. Reuse, reduce, support: Design principles for green data mining. **Business & Information Systems Engineering**, Springer, 2023.

SIEGLE, D. A role for chatgpt and ai in gifted education. **Gifted Child Today**, SAGE Publications Sage CA: Los Angeles, CA, 2023.

SPATOLA, N.; URBANSKA, K. God-like robots: the semantic overlap between representation of divine and artificial entities. **Ai & Society**, Springer, v. 35, n. 2, p. 329–341, 2020.

STROPARO, T. R.; ARAÚJO, J. H. K. de; BORTOLOTTI, M. A.; JUNIOR, O. d. S. L. Inteligência artificial na gestão de custos: avanços, desafios e oportunidades. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 10, n. 6, p. 1446–1456, 2024.

SUN, S.; ZHAI, Y.; SHEN, B.; CHEN, Y. Newspaper coverage of artificial intelligence: A perspective of emerging technologies. **Telematics and Informatics**, Elsevier, 2020.

WIRTH, R.; HIPPEL, J. Crisp-dm: Towards a standard process model for data mining. In: MANCHESTER. **Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining**. [S.l.], 2000. v. 1, p. 29–39.

APÊNDICE A – ARTIGO ORIGINAL

Artificial Intelligence in the Spotlight: A Study of Brazilian News Coverage

Jonathan Oliveira Fernandez¹, Andrey Silva Pontes¹,
Solange Oliveira Rezende² Fábio Manoel França Lobato¹

¹Universidade Federal do Oeste do Pará (UFOPA)
Santarém - PA - Brasil

²Universidade de São Paulo (USP)
São Carlos - SP - Brasil.

fabio.lobato@ufopa.edu.br

Abstract. *Artificial Intelligence (AI) has become increasingly integrated into different layers of society, bringing both benefits and risks. With the growth of this technology, media coverage of it has also intensified. This paper analyzes media coverage of AI in Brazil, identifying the predominant themes and trends in news about the technology. To this end, four news portals were analyzed using topic modeling. The results show that the news addresses the impact of AI on the job market, advances in health, political issues involving regulation and disinformation, and technological innovations. Therefore, the study provides a comprehensive view of the topics highlighted by the media on the main Brazilian portals, contributing to the state of the art by promoting insights that can serve as a basis for the formulation of public policies and regulations regarding the technology.*

Resumo. *A Inteligência Artificial (IA) tem se integrado cada vez mais em diferentes camadas da sociedade, trazendo benefícios e riscos. Com o crescimento dessa tecnologia, a cobertura midiática sobre o tema também se intensificou. Assim, este trabalho analisa a cobertura midiática da IA no Brasil, identificando os temas e tendências predominantes nas notícias relacionadas à tecnologia. Para isso, foram analisados quatro portais jornalísticos utilizando modelagem de tópicos. Os resultados obtidos revelam que as notícias abordam o impacto da IA no mercado de trabalho, os avanços na saúde, as questões políticas relacionadas à regulamentação e desinformação, além das inovações tecnológicas. Portanto, o estudo fornece uma visão abrangente dos temas destacados pela mídia nos principais portais brasileiros, contribuindo para o estado da arte ao promover insights que podem servir de subsídio para a formulação de políticas públicas e regulamentações relacionadas à tecnologia.*

1. Introdução

O avanço tecnológico tem contribuído para simplificar cada vez mais as tarefas humanas, reduzindo a necessidade de esforço manual em várias atividades e, em alguns casos, até eliminando a necessidade de realizá-las, graças ao uso de algoritmos de Inteligência Artificial (IA) por empresas, governos e organizações [Ntoutsis et al. 2020]. O uso dessa tecnologia oferece diversas vantagens que favorecem a tomada de decisões, uma vez que

a capacidade de analisar grandes volumes de dados de forma rápida pode ser benéfica para a otimização de processos, a identificação de tendências e o aprimoramento da eficiência operacional [Stroparo et al. 2024].

Um dos avanços mais relevantes na área nos últimos anos tem sido o desenvolvimento e uso dos *Large Language Models* (LLMs), amplamente conhecidos como IA generativa, nos quais os algoritmos são capazes de produzir conteúdo textual, visual e auditivo de maneira quase autônoma [Longoni et al. 2022]. Essa digitalização dos meios de comunicação impactou questões que antes eram consideradas exclusivamente humanas, uma vez que os algoritmos estão auxiliando os humanos em tarefas como a redação de textos, e-mails e livros [Peña-Fernández et al. 2023]. No entanto, apesar do grande potencial da IA generativa, sua implementação ainda enfrenta uma série de desafios, como a discriminação racial na geração de imagens¹, em que os modelos de IA podem reproduzir e amplificar preconceitos existentes se forem treinados com dados que contêm viés [Garcia 2020]. Além disso, a técnica *deepfake*, que, embora traga benefícios, pode ser utilizada para gerar desinformação tanto no âmbito social quanto político [Molina and Berenguel 2022].

Desse modo, um objetivo crítico para o desenvolvimento da IA é a ampliação da confiança pública, uma vez que a população desempenha um papel fundamental na formulação de políticas públicas e na supervisão regulatória dessa tecnologia. A opinião popular pode influenciar tanto o tipo de IA que será desenvolvida quanto a forma como ela será regulamentada pelo governo [Ouchchy et al. 2020]. Assim, à medida que a IA generativa avança, um número crescente de notícias surge, explorando como os efeitos dessa evolução impactam nossas vidas [Choi 2024]. Nesse sentido, a mídia jornalística desempenha um papel preponderante ao traduzir conhecimentos complexos em informações acessíveis, e durante esse processo, influencia a percepção pública sobre a tecnologia, especialmente quando as pessoas se sentem desconfiadas ou ansiosas [Chuan 2023].

Considerando que as percepções do público sobre o tema da IA são moldadas pela maneira como a mídia apresenta essas informações, o objetivo deste trabalho é analisar a cobertura midiática a respeito desse tema e entender como os principais portais jornalísticos brasileiros relatam as notícias, quais temas são predominantes e de que maneira essa narrativa pode impactar a opinião pública sobre a tecnologia. Para essa finalidade, foram utilizadas técnicas de Processamento de Linguagem Natural (PLN) e modelagem de tópicos com a técnica BERTopic, visando identificar padrões temáticos recorrentes nas notícias e compreender como a mídia constrói e dissemina essas narrativas. Ao realizar essa análise, é possível observar quais assuntos recebem maior destaque e os possíveis impactos éticos, econômicos e tecnológicos, contribuindo assim para o debate sobre a regulação dessa tecnologia emergente.

2. Trabalhos Relacionados

No trabalho [Ouchchy et al. 2020], os autores analisaram a cobertura midiática da IA com enfoque nas questões éticas relacionadas à tecnologia. Utilizando o banco de dados *NexisUni*, os pesquisadores coletaram os dados no período de 2014 a 2018, resultando em uma amostra de 254 publicações, incluindo notícias, artigos de revistas e blogs. As publicações foram categorizadas em cinco áreas: questões éticas; princípios baseados em estruturas

¹Disponível em: <http://glo.bo/3M9fsat>

éticas; recomendações das publicações; tom da publicação; tipo de tecnologia específica. Os resultados revelaram que inicialmente as reportagens possuíam um tom otimista e entusiasmado, que se tornou mais crítico ou neutro ao longo dos anos. Os temas éticos mais discutidos foram preconceito e proteção de dados, enquanto as recomendações ainda estão em estágios iniciais. O trabalho sugere que a mídia mantém uma abordagem prática na cobertura da ética da IA, mas muitos autores carecem de conhecimento adequado sobre o tema. Em contraste com o presente estudo, o trabalho correlato não menciona a utilização de métodos automatizados para análise dos dados (*e.g.* modelagem de tópicos), sendo este um diferencial deste trabalho.

O estudo de [Cools et al. 2024] examinou o panorama midiático sobre IA e automação em jornais dos Estados Unidos, analisando a forma que a mídia apresenta os benefícios e riscos associados a estes temas. Por meio da *Application Programming Interface* (API) dos portais *New York Times* (NYT) e *The Washington Post* (WP), os pesquisadores coletaram notícias de 1985 a 2020. Posteriormente, realizaram modelagem de tópicos com *Latent Dirichlet Allocation* (LDA) e análise manual de enquadramento para a análise dos dados, na qual rotularam as notícias como neutras, positivas e negativas em uma escala de utópico a distópico. Os resultados revelaram dez tópicos e quadros distintos. Os tópicos abordaram temas como Política, Educação, e Esporte, com oito tópicos possuindo relevância até um determinado período, e dois que permanecem relevantes ao longo do tempo, nos quais observou-se um aumento a partir de 2013 nos tópicos Trabalho, Arte e Educação. Em relação aos enquadramentos, a mídia tem sido mais otimista do que pessimista ao longo do tempo, embora ainda existam narrativas distópicas que retratam a IA e a automação como tecnologias que irão substituir os humanos no futuro. Portanto, o estudo apresenta uma visão geral de longo prazo sobre estas tecnologias, revelando tanto temas positivos, quanto negativos. Apesar do estudo correlato também utilizar modelagem de tópicos, o presente trabalho se diferencia ao empregar a técnica BERTopic, que permite uma análise mais detalhada ao incorporar informações contextuais na formação dos tópicos, cuja eficácia foi comprovada por [Egger and Yu 2022].

O trabalho de [Sun et al. 2020] realizou uma análise na cobertura midiática sobre IA ao longo das últimas quatro décadas, utilizando como fontes os veículos de notícias NYT, WP, *The Guardian* e *USA Today*. Os autores analisaram 1.776 notícias usando *K-Means* para identificar tópicos e temas principais, análise de léxico para detectar padrões de argumentação, análise de rede para mapear países e organizações mencionadas e análise de correspondência múltipla para explorar associações entre os dados. Os resultados revelaram 14 *clusters*, dos quais os pesquisadores extraíram as palavras-chave mais relevantes de cada tema, revelando tópicos como “Assistência médica”, “Emprego e economia”, “Regulamentação e política”, entre outros. Na análise de léxico identificaram nove padrões de argumentação, no qual o mais prevalente foi o pragmático, relacionado a notícias sobre aplicação prática da IA no cotidiano. A análise de rede destacou potências como Estados Unidos, China, Rússia, Japão e Alemanha, além das organizações mais citadas como Google, Facebook e Microsoft. Por fim, a análise de correspondência múltipla revelou que o NYT tende a focar mais em notícias sobre negócios, enquanto o *The Guardian* se concentra em assuntos sobre robôs e algoritmos; os demais portais estão relacionados às notícias sobre aplicações da IA. Assim, o estudo conclui que embora a mídia aborde IA há décadas, a cobertura se intensificou somente nos últimos cinco anos, adotando uma abordagem mista e equilibrada.

3. Metodologia

Seguindo o guia proposto por [Wirth and Hipp 2000], este estudo utilizou a metodologia *Cross-Industry Standard Process for Data Mining* (CRISP-DM), amplamente reconhecida na literatura para conduzir análises de dados e mineração de textos, conforme demonstrado em estudos como [Schneider et al. 2023, Costa et al. 2022]. O CRISP-DM é composto por seis etapas: Entendimento do negócio, Entendimento dos dados, Preparação dos dados, Modelagem, Avaliação e Implantação; detalhadas nas subseções a seguir.

3.1. Entendimento do negócio

A análise de notícias sobre IA foi enriquecida pela experiência das pessoas autoras e reuniões com especialistas da área, que incluíram debates com membros do Grupo de Trabalho em Economia Digital do G20, participações em eventos como o *AI Summit* e colaborações com acadêmicos e profissionais de diferentes setores. A combinação da *expertise* das pessoas autoras e do conhecimento obtido com outros especialistas possibilitou uma compreensão mais profunda sobre as implicações sociais, econômicas e éticas da IA, proporcionando a base necessária para este trabalho.

Paralelamente foi realizado um levantamento da literatura do tipo *Ad Hoc* [Ralph and Baltes 2022], visando compreender a importância da análise de notícias, uma vez que a forma como essa tecnologia é retratada pode afetar sua aceitação e a formação de políticas públicas, no qual os principais estudos encontrados foram apresentados tanto na Seção 1 quanto na Seção 2.

3.2. Entendimento dos dados

Esta etapa foi dividida em duas fases: seleção de palavras-chave e coleta de dados. Para a primeira fase foi utilizado o método *Delphi* apresentado no estudo de [Baines and Regan de Bere 2018], que consiste na concordância entre especialistas, onde as palavras-chave são listadas e validadas, resultando nas seguintes palavras-chave: “Inteligência Artificial”, “OpenAI”, “Google Gemini” e “Microsoft Copilot”.

Na fase de coleta de dados os portais jornalísticos selecionados foram o Portal G1, Metrôpoles, Folha de São Paulo, e CNN Brasil devido à popularidade e abrangência na cobertura de notícias destes portais, bem como a diversidade editorial e a credibilidade consolidada no jornalismo brasileiro. Os dados foram coletados por meio de *web scraping* utilizando um *web crawler* e, posteriormente, salvos em formato *Comma-Separated Values* (CSV) para facilitar a manipulação e análise eficiente. Os atributos coletados incluem: 1) título da notícia; 2) resumo; 3) corpo do texto; 4) data da publicação; 5) região; 6) link da notícia; 7) palavra-chave; e 8) plataforma de notícia.

3.3. Preparação dos dados

Esta etapa também foi dividida em duas fases: filtragem e pré-processamento. Utilizando como base a teoria fundamentada descrita por [Baumer et al. 2017], foi realizada uma filtragem para identificar quais notícias se enquadravam no escopo do estudo, já que muitas delas apenas mencionavam IA de passagem, sem discutir o tema de maneira substancial. O processo envolveu a leitura dos títulos e conteúdos das notícias, focando na relação da notícia com a palavra-chave, a fim de garantir a consistência dos dados e mitigar possíveis

inconsistências. As notícias que não abordavam diretamente o assunto foram removidas do conjunto de dados.

Em seguida, os textos passaram para a fase de pré-processamento, visando a remoção de ruídos, palavras irrelevantes, pontuações e acentos que podem afetar a qualidade e relevância das informações extraídas. Para isso foram aplicados *scripts* escritos em *Python* seguindo o *pipeline* de [Araújo et al. 2024], que envolve os seguintes passos: i) padronização para caixa-baixa; ii) remoção de pontuação, numerais e caracteres especiais; iii) remoção de *stopwords*².

3.4. Modelagem

Nesta etapa, foram conduzidas análises com o objetivo de identificar temas recorrentes nas notícias, utilizando técnicas de PLN e modelagem de tópicos com o BERTopic. Essa técnica cria tópicos a partir de *embeddings* gerados por modelos de linguagem pré-treinados. O BERTopic gera *embeddings* contextuais para os textos, capturando o significado semântico das palavras em diferentes contextos. Em seguida, realiza a clusterização dos textos que possuem semântica semelhante, e os tópicos são extraídos utilizando a medida *Class-Based Term Frequency-Inverse Document Frequency* (c-TF-IDF), que considera a frequência e o contexto dos termos no corpus do documento [Grootendorst 2022].

Para a análise, foi empregada a configuração padrão do BERTopic, utilizando o modelo *Multilingual*, que possui suporte para a língua portuguesa. A fim de proporcionar uma visão geral e detalhada, optou-se por segmentar os tópicos por portal jornalístico, selecionando os temas mais distintos entre eles. Posteriormente, os tópicos gerados foram anotados manualmente para uma melhor interpretação dos resultados.

3.5. Avaliação e Implantação

A etapa de Avaliação manteve a abordagem da teoria fundamentada, no qual uma amostra representativa dos dados foi separada com base em um intervalo de confiança de 95% e uma taxa de erro de 10%, para que os autores realizassem a leitura dessa amostra e em seguida confrontassem com os resultados obtidos pelos algoritmos, objetivando a detecção de incongruências ou informações incorretas que pudessem ter passado despercebidas nas etapas anteriores. Posteriormente esses resultados foram apresentados ao grupo de pesquisa e avaliados em conjunto utilizando a técnica *ground truth*, além da discussão com os especialistas do domínio que participaram do entendimento do negócio.

A etapa de implantação é apoiada por um relatório técnico contendo os resultados obtidos ao longo do estudo, visando proporcionar *insights* relevantes para subsidiar o planejamento de políticas públicas.

4. Resultados

Nesta seção são apresentados e discutidos os resultados obtidos considerando a metodologia descrita na seção anterior. Primeiramente são apresentados os resultados referentes à coleta de dados, seguidos pelos resultados obtidos com a modelagem de tópicos.

²Palavras que não têm relevância, mas aparecem com frequência (e.g. preposições, artigos, pronomes).

4.1. Coleta de dados

Foram coletadas 12.308 notícias publicadas nos portais jornalísticos no período de 2016 a 2024. Após a aplicação do pré-processamento e da filtragem, restaram 2.557 notícias, as quais foram utilizadas para análises posteriores. Por meio da análise exploratória do conjunto de dados foi possível obter a distribuição da quantidade de notícias por plataforma, conforme apresentado na Figura 1.

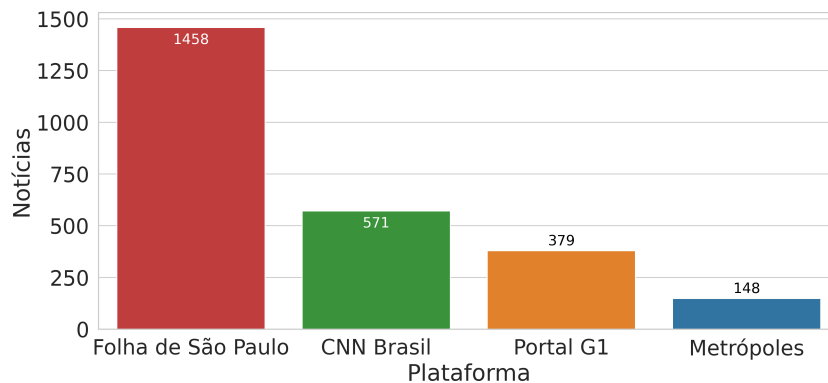


Figura 1. Quantidade de notícias por plataforma.

Vide a Figura 1 é possível observar que o portal Folha de São Paulo possui o maior número de notícias, o que pode ser atribuído ao fator de que o jornal é um dos maiores veículos de comunicação no Brasil, onde possui características como textos de qualidade e conteúdos diferentes dos tabloides, no qual priorizam tanto a cobertura da mídia nacional quanto a local (*i.e.* São Paulo) [Marques et al. 2020]. Além dos resultados mencionados, o volume de publicações anuais nas plataformas também apresentou variações ao longo dos anos analisados, refletindo mudanças no cenário jornalístico sobre IA. A Figura 2 apresenta a distribuição anual de notícias publicadas durante o período analisado.

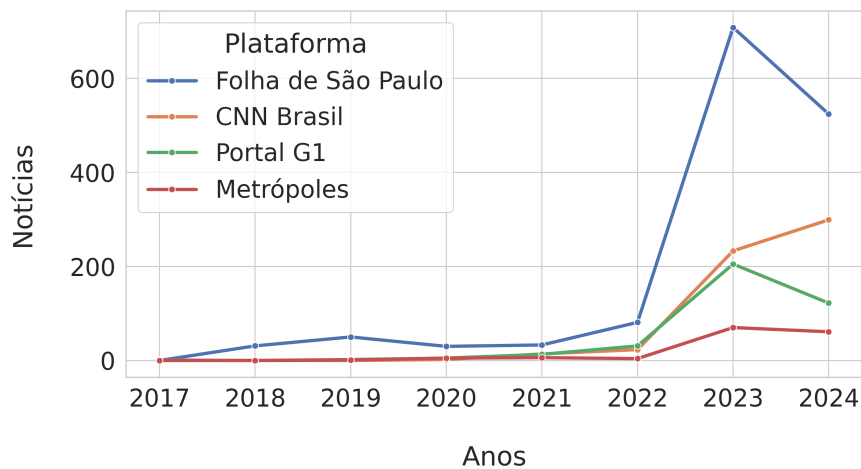


Figura 2. Quantidade de notícias por ano.

Tabela 1. Tópicos e termos das notícias por plataforma.

Portal	Tópicos	Termos
Folha de São Paulo	Transformação no trabalho	trabalho, empregos, inteligencia_artificial, empresa, trabalhadores, funcoes, negocios, anos, profissionais, contratacoes
	Futuro da saúde	pacientes, cancer, medicos, diagnostico, doenca, saude, dados, algoritmos, paciente, doencas
CNN Brasil	Mercado tecnológico	openai, microsoft, inteligencia_artificial, empresa, altman, amazon, chatgpt, tecnologia, ceo, demissao
	Dispositivos inteligentes	apple, iphone, inteligencia_artificial, ipad, intelligence, empresa, google, recursos, siri, dispositivos
Portal G1	Política e IA	inteligencia_artificial, eleitoral, tse, eleicoes, uso, senado, proposta, presidente, regras, texto
	Regulamentação da IA	européia, uniao, dados, inteligencia_artificial, regras, acordo, uso, chatgpt, europeu, empresa
Metrópoles	Manipulação digital	video, deepfake, lula, inteligencia_artificial, fake, conteudo, uso, desinformacao, tse, tecnica
	Chats inteligentes	chatgpt, openai, inteligencia_artificial, ferramenta, chatbot, bard, usuarios, empresa, robo, informacoes

termos indicam que este tópico está ligado a notícias sobre o impacto da IA nos empregos, abordando como esta tecnologia está moldando o futuro do trabalho, o que inclui discussões sobre a automação de tarefas, mudanças nas vagas de emprego e substituição de funções realizadas por humanos. [George et al. 2023] afirmam que o impacto varia entre os setores, com alguns experimentando transformações mais profundas que outros. Por exemplo, o setor de finanças foi altamente impactado pela automação de tarefas como detecção de fraudes e análise de riscos.

O setor de saúde tem apresentado avanços promissores na melhoria de diagnósticos e atendimentos ao paciente, evidenciando o potencial transformador da IA em diferentes áreas. O que está ligado ao tópico “Futuro da saúde” que aborda notícias sobre os avanços da IA na saúde e bem-estar da sociedade, abrangendo as inovações tecnológicas que impulsionam este setor, como o monitoramento de pacientes gestantes³. No entanto, [Bunz and Braghieri 2022] informam que as manchetes devem ser cuidadosas ao noticiar as contribuições da IA para a saúde, por haver o risco de descrever os sistemas de IA como algo superior ao humano, antropomorfizando a tecnologia. Esse cuidado é reforçado por estudos como o de [Spatola and Urbanska 2020], que identifica uma tendência crescente em atribuir expectativas positivas, e até mesmo divinas na IA, o que pode afetar a percepção do público sobre o papel dessa tecnologia.

O portal CNN Brasil apresenta destaque na cobertura sobre a área mercadológica da IA, apresentando tópicos sobre o atual mercado e a introdução de novos recursos inteligentes nos dispositivos. No tópico “Mercado tecnológico” é possível observar por meio dos termos a cobertura sobre as grandes empresas como OpenAI e Microsoft, explorando suas inovações e estratégias de mercado, visto que estas empresas são parceiras de mercado⁴. As notícias reportam como as empresas estão integrando recursos inteligentes em seus produtos e serviços⁵, um exemplo notável destas inovações é o uso da tecnologia

³Disponível em: <https://bit.ly/3Ai3G7>

⁴Disponível em: <https://bit.ly/microsoft-partnerships>

⁵Disponível em: <https://bit.ly/microsoft-openai>

do ChatGPT no mecanismo de busca do navegador Edge, no qual a IA é utilizada para oferecer melhores resultados de busca [Siegler 2023].

O tópico “Dispositivos inteligentes” também indica uma cobertura no lançamento de dispositivos e tecnologias baseadas em IA, como a *Apple Intelligence*, um novo modelo de IA que está sendo desenvolvido para aprimorar a experiência do usuário em dispositivos como *iPhone*, *iPad* e *Mac*, incluindo funcionalidades para escrita, geração de imagens e integrações com a assistente virtual Siri [Mohamed and Lucke-Wold 2024]. Além disso, a cobertura inclui inovações do Google, como os chips *Tensor Processing Unit* (TPU) projetados para otimizar o desempenho da IA nas tarefas de aprendizado de máquina, PLN e entre outros. Notavelmente, a Apple tem integrado esses chips TPUs em seus processos de treinamento de modelos de IA⁶, refletindo o impacto dos desenvolvimentos inovadores do Google no campo de IA.

A cobertura do portal G1 está inclinada para um aspecto político, abordando temas como eleições e a regulamentação da IA na Europa. O tópico “Política e IA”, foca em como a IA está influenciando os processos eleitorais, desde o uso para as campanhas até a disseminação de desinformação por meio de *deepfakes* nas redes sociais⁷, uma vez que os conteúdos manipulados podem ser utilizados para enganar eleitores e influenciar a opinião pública. De acordo com [Molina and Berenguel 2022], o uso da tecnologia para fins políticos gera preocupações, podendo ser uma ameaça à democracia. Assim, o Brasil tem avançado na formulação de políticas para lidar com os impactos da IA, com a criação do Projeto de Lei (PL) que visa regulamentar o desenvolvimento e o uso da IA (PL 2.338/2023), refletindo um esforço para garantir que a tecnologia seja usada de forma ética e segura.

Assim como na União Europeia (UE), que foi a região pioneira na regulamentação da IA. O tópico “Regulamentação da IA” está associado as notícias que tratam sobre como a UE visa moldar a regulamentação global de tecnologia, replicando o sucesso do Regulamento Geral de Proteção de Dados (GDPR), considerado um modelo de alta qualidade para as leis de proteção de dados, no qual vários países basearam suas estruturas de leis nesse modelo [Belli et al. 2023].

No portal Metrôpoles os tópicos estão associados à manipulação de dados e assistentes virtuais, como *chatbots*. O tópico “Manipulação digital” está relacionado às notícias sobre como os dados são processados e utilizados para a disseminação de desinformação, desde notícias falsas até *deepfakes*, onde muitas dessas práticas envolvem celebridades, políticos⁸ e diretores executivos de empresas, uma vez que o fácil acesso de fotos e vídeos na internet facilitam a criação e circulação desses conteúdos [Molina and Berenguel 2022]. Além disso, estas práticas também podem ser utilizadas para causar danos a outras pessoas, ocasionando em consequências graves como a difamação e invasão de privacidade.

O tópico “Manipulação Digital” do portal Metrôpoles apresenta similaridade com o tópico “Política e IA” abordado pelo portal G1. Ambos os tópicos destacam como a manipulação de dados e a propagação de desinformação são temas preocupantes no

⁶Disponível em: <https://bit.ly/apple-chips-google>

⁷Disponível em: <http://glo.bo/3Mnm4Cj>

⁸Disponível em: <https://bit.ly/deepfake-politicos>

cenário atual, ilustrando as consequências dessas práticas e ressaltando a necessidade de impôr regras éticas para o uso e desenvolvimento dessas tecnologias, visando mitigar os impactos negativos na sociedade.

Por fim, o tópico “Chats inteligentes” explora o desenvolvimento e aplicação de *chatbots* e outras tecnologias de IA. [Maedche et al. 2019] apontam que os assistentes virtuais estão se tornando cada vez mais comuns e integrados ao cotidiano dos usuários, por exemplo, assistentes que utilizam comando de voz como a Amazon Alexa ou interfaces baseadas em texto para o atendimento ao cliente. Ainda, os autores apresentam preocupações em torno dessas tecnologias, gerando debates sobre fraudes e enganosa, visto que elas apresentam semelhanças com a linguagem humana.

Os tópicos apresentados e segmentados por portais de notícias revelam a amplitude da cobertura jornalística sobre a IA, apresentando as transformações em várias áreas, os desafios políticos e questões sobre a ética e regulação da tecnologia. Essa análise demonstra o acompanhamento da mídia sobre a evolução da IA, refletindo o papel central da imprensa em conscientizar o público sobre os benefícios e riscos associados a tecnologia, contribuindo para a percepção do público sobre a IA.

5. Conclusões

Os avanços da IA têm trazido inúmeras facilidades e benefícios para a sociedade, mas também apresentam desafios e riscos significativos onde a falta de ética, supervisão regulatória adequada e responsabilidade podem resultar em consequências maiores. Nesse sentido, a mídia jornalística atua como uma ferramenta de divulgação e discussão sobre esses desafios, promovendo uma compreensão crítica da população a respeito do progresso de tecnologias emergentes como a IA.

Nesse contexto, o presente trabalho teve como objetivo analisar a cobertura midiática da IA no Brasil, investigando como os diferentes portais de notícias abordam a IA e suas implicações. Foram analisadas 2.557 notícias por meio de uma modelagem de tópicos, no qual os resultados revelaram temas como mercado de trabalho, saúde, política, manipulação de conteúdo e inovações na área. O estudo fornece uma visão geral de quais temas estão sendo destacados na mídia, oferecendo contribuições para o estado da arte ao apresentar como a IA é retratada; fornecendo subsídios para debates sobre os desafios éticos, sociais e econômicos advindos dessa tecnologia.

Para trabalhos futuros, pretende-se aumentar a quantidade de análises no conjunto de dados, como análise de sentimentos ou rede semânticas, para oferecer uma perspectiva mais detalhada. Ademais, a utilização de LLMs como ferramenta de análise pode permitir uma interpretação mais rica das narrativas sobre a IA na mídia jornalística, identificando nuances linguísticas e contextuais. Outra sugestão de trabalho futuro é explorar a etapa de implantação do CRISP-DM, com um relatório técnico contendo os resultados obtidos ao longo do estudo, a fim de proporcionar *insights* relevantes para subsídio de políticas públicas, como a regulação da IA, de forma que recurso seja conduzido de maneira responsável, visando o bem-estar da sociedade como um todo.

Agradecimentos

Este trabalho foi apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) DT-303031/2023-9, PQ-309575/2021-4, PIBIC - 166028/2024-0; e

pela Fundação Amazônia de Amparo a Estudos e Pesquisas (FAPESPA) PRONEM-FAPESPA/CNPq nº 045/2021.

Referências

- Araújo, G. D. S., Fernandez, J. O., Da Silva, M. S., and Lobato, F. M. F. (2024). Persona and issue analysis on tourism social media: a case study of pirenópolis, goiás, brazil. In *Proceedings of the 20th Brazilian Symposium on Information Systems*, pages 1–10.
- Baines, R. L. and Regan de Bere, S. (2018). Optimizing patient and public involvement (ppi): Identifying its “essential” and “desirable” principles using a systematic review and modified delphi methodology. *Health Expectations*, 21(1):327–335.
- Baumer, E. P., Mimno, D., Guha, S., Quan, E., and Gay, G. K. (2017). Comparing grounded theory and topic modeling: Extreme divergence or unlikely convergence? *Journal of the Association for Information Science and Technology*, 68(6):1397–1410.
- Belli, L., Curzi, Y., and Gaspar, W. B. (2023). Ai regulation in brazil: Advancements, flows, and need to learn from the data protection experience. *Computer Law & Security Review*, 48:105767.
- Bunz, M. and Braghieri, M. (2022). The ai doctor will see you now: assessing the framing of ai in news coverage. *AI & society*, 37(1):9–22.
- Choi, S. (2024). Temporal framing in balanced news coverage of artificial intelligence and public attitudes. *Mass Communication and Society*, 27(2):384–405.
- Chuan, C.-H. (2023). A critical review of news framing of artificial intelligence. *Handbook of Critical Studies of Artificial Intelligence*, pages 266–276.
- Cools, H., Van Gorp, B., and Opgenhaffen, M. (2024). Where exactly between utopia and dystopia? a framing analysis of ai and automation in us newspapers. *Journalism*.
- Costa, G. d. S., Couto, D. C., Junior, A. F. J., and Lobato, F. M. (2022). Feminismo e redes sociais online: uma análise de tweets sobre o dia internacional da mulher. In *Anais do XI Brazilian Workshop on Social Network Analysis and Mining*, pages 169–180. SBC.
- Egger, R. and Yu, J. (2022). A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in sociology*, 7:886498.
- Garcia, A. C. (2020). Ética e inteligencia artificial. *Computação Brasil*, (43):14–22.
- George, A. S., George, A. H., and Martin, A. G. (2023). Chatgpt and the future of work: a comprehensive analysis of ai’s impact on jobs and employment. *Partners Universal International Innovation Journal*, 1(3):154–186.
- Grootendorst, M. (2022). Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.
- Longoni, C., Fradkin, A., Cian, L., and Pennycook, G. (2022). News from generative artificial intelligence is believed less. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pages 97–106.
- Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., and Söllner, M. (2019). Ai-based digital assistants: Opportunities, threats, and research perspectives. *Business & Information Systems Engineering*, 61:535–544.

- Marques, F. P. J., Miola, E., Mitozo, I., and Mont'Alverne, C. (2020). Similar, but not the same: Comparing editorial and news agendas in brazilian newspapers. *Journalism Practice*, 14(9):1066–1086.
- Mohamed, A. A. and Lucke-Wold, B. (2024). Apple intelligence in neurosurgery. *Neurosurgical Review*, 47(1):327.
- Molina, A. C. and Berenguel, O. L. (2022). Deepfake: A evolução das fake news. *Research, Society and Development*, 11(6):e56211629533–e56211629533.
- Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdil, W., Vidal, M.-E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., et al. (2020). Bias in data-driven artificial intelligence systems—an introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3):e1356.
- Ouchchy, L., Coin, A., and Dubljević, V. (2020). Ai in the headlines: the portrayal of the ethical issues of artificial intelligence in the media. *AI & SOCIETY*, 35:927–936.
- Peña-Fernández, S., Meso-Ayerdi, K., Larrondo-Ureta, A., and Díaz-Noci, J. (2023). Without journalists, there is no journalism: the social dimension of generative artificial intelligence in the media. *El Profesional de la información*, 32(2).
- Ralph, P. and Baltes, S. (2022). Paving the way for mature secondary research: the seven types of literature review. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, pages 1632–1636.
- Schneider, J., Seidel, S., Basalla, M., and vom Brocke, J. (2023). Reuse, reduce, support: Design principles for green data mining. *Business & Information Systems Engineering*.
- Siegle, D. (2023). A role for chatgpt and ai in gifted education. *Gifted Child Today*.
- Spatola, N. and Urbanska, K. (2020). God-like robots: the semantic overlap between representation of divine and artificial entities. *Ai & Society*, 35(2):329–341.
- Stroparo, T. R., de Araújo, J. H. K., Bortolotti, M. A., and Junior, O. d. S. L. (2024). Inteligência artificial na gestão de custos: avanços, desafios e oportunidades. *Revista Ibero-Americana de Humanidades, Ciências e Educação*, 10(6):1446–1456.
- Sun, S., Zhai, Y., Shen, B., and Chen, Y. (2020). Newspaper coverage of artificial intelligence: A perspective of emerging technologies. *Telematics and Informatics*.
- Wirth, R. and Hipp, J. (2000). Crisp-dm: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, volume 1, pages 29–39. Manchester.

APÊNDICE B – TÓPICOS GERAIS

Tabela B.1 – Tópicos e termos das notícias.

Tópicos	Termos
Inovação Tecnológica	inteligencia_artificial, empresa, chatgpt, tecnologia, google, dados, uso, microsoft, ferramenta, openai
IA e mídias sociais	video, inteligencia_artificial, imagens, deepfake, musica, pessoas, foto, imagem, sociais, tecnologia
Liderança da OpenAI	openai, elon_musk, altman, empresa, inteligencia_artificial, sam_altman, ceo, chatgpt, demissao, sutskever
Regulamentação da IA	européia, inteligencia_artificial, ue, uniao, lei, europeu, ucrania, russia, uso, parlamento
Dispositivos inteligentes	apple, iphone, inteligencia_artificial, empresa, ipad, intelligence, pro, iphones, dispositivos, google
Mercado tecnológico	importa, recebe, mercado, inteligencia_artificial, petrobras, nvidia, chatgpt, movido, elon_musk, startup
Restauração de documentos históricos	pergaminho, vesuvio, papiro, decifrar, pesquisadores, carbonizado, erupcao, tinta, textos, herculano
Plataformas da Meta	whatsapp, meta, inteligencia_artificial, consultas, instagram, facebook, usuarios, chatgpt, campinas, faltas
IA para áudios e vozes	voz, sky, openai, vozes, atriz, scarlett_johansson, chatgpt, johansson, empresa, engine
Proteção de dados	dados, italia, agencia, chatgpt, protecao, openai, chatbot, investigacao, violacao, usuarios

APÊNDICE C – TÓPICOS SEGMENTADOS POR PERÍODOS ANUAIS

Tabela C.1 – Tópicos e termos das notícias (2016 a 2021).

Tópicos	Termos
IA e suas aplicações	inteligencia_artificial, humanos, robo, tecnologia, ha, planetas, pesquisadores,xadrez, jogo, algoritmo
Uso de chatbots	inteligencia_artificial, brasil, tecnologia, atendimento, uso, desenvolvimento, paciente, milhoes, sistemas, texto
Pesquisas em IA	google, empresa, inteligencia_artificial, gebruik, funcionarios, companhia, equipe, pesquisa, hieroglifos, trabalho
Aplicações da IA na saúde	cancer, pacientes, inteligencia_artificial, saude, diagnosticos, diagnostico, imagem, dados, tratamento, exame
Impacto dos deepfakes	video, deepfakes, cruise, ume, deepfake, falsos, imagens, inteligencia_artificial, conteudo, twitter

Tabela C.2 – Tópicos e termos das notícias (2022 a 2024).

Tópicos	Termos
Avanços da IA	inteligencia_artificial, empresa, tecnologia, google, chatgpt, microsoft, uso, dados, estados_unidos, brasil
IA para criação de conteúdos falsos	inteligencia_artificial, video, imagens, deepfake, musica, foto, pessoas, sociais, imagem, conteudo
Interação com IA por voz	chatgpt, openai, inteligencia_artificial, usuarios, ferramenta, empresa, chatbot, voz, gpt, robo
Liderança da OpenAI	openai, elon_musk, altman, empresa, inteligencia_artificial, sam_altman, ceo, chatgpt, demissao, microsoft
Futuro da saúde	pacientes, diagnostico, cancer, doenca, estudo, pesquisadores, inteligencia_artificial, exames, saude, parkinson
Regulamentação da IA	européia, inteligencia_artificial, ue, uniao, lei, europeu, russia, ucrania, parlamento, regras
Dispositivos inteligentes	apple, iphone, inteligencia_artificial, empresa, ipad, intelligence, pro, dispositivos, iphones, google
Mercado tecnológico	importa, recebe, mercado, inteligencia_artificial, chatgpt, petrobras, demissoes, nvidia, movido, elon_musk
Inovações da Amazon em IA	amazon, amazonia, alexa, desmatamento, inteligencia_artificial, empresa, anthropic, ano, generativa, tecnologia
Proteção de dados	dados, italia, agencia, chatgpt, protecao, openai, chatbot, investigacao, violacao, usuarios

ANEXO A – REGISTRO DE PARTICIPAÇÃO NO EVENTO

CERTIFICADO

A Sociedade Brasileira de Computação atesta a participação de

Jonathan Oliveira Fernandez

no(a) XXXIV Brazilian Conference on Intelligent Systems (BRACIS 2024),

realizado(a) entre 17 e 21 de Novembro em Belém - PA, com carga horária de 40 horas.

Prof. Thais Vasconcelos Batista
Presidente da Sociedade Brasileira de Computação



Prof. Denis Lima do Rosário
Diretor de Eventos e Comissões Especiais