

Estado da publicação: Não informado pelo autor submissor

NARRATIVAS DIGITAIS E INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE COMPARATIVA TRIDIMENSIONAL DE CONTEÚDOS AUDIOVISUAIS

Allan Herison Ferreira, Ana Carolina Trevisan

<https://doi.org/10.1590/SciELOPreprints.9937>

Submetido em: 2024-09-12

Postado em: 2024-09-14 (versão 1)

(AAAA-MM-DD)

A moderação deste preprint recebeu o endosso de:

Carla Baptista (ORCID: <https://orcid.org/0000-0002-8188-3567>)

NARRATIVAS DIGITAIS E INTELIGÊNCIA ARTIFICIAL: UMA ANÁLISE COMPARATIVA TRIDIMENSIONAL DE CONTEÚDOS AUDIOVISUAIS

ALLAN HERISON FERREIRA/A 1

ORCID: <https://orcid.org/0000-0003-3606-2089>

allanherison@gmail.com.br

ICNOVA/Universidade NOVA de Lisboa. Lisboa, Portugal | LAPS/Universidade de São Paulo, SP, Brasil.

ANA CAROLINA TREVISAN/A 2

ORCID: <https://orcid.org/0000-0002-2818-5858>

anacarolinatcf@gmail.com.br

IFILNOVA/Universidade NOVA de Lisboa. Lisboa, Portugal

RESUMO: Este artigo examina o impacto das plataformas de redes sociais e das tecnologias de inteligência artificial (IA) na construção de identidades digitais e na comunicação. Ele explora como plataformas como YouTube influenciam a criação, expressão e percepção de identidades, ao mesmo tempo em que discute o papel dos algoritmos de IA na modulação dessas interações. A pesquisa adota uma abordagem comparativa tridimensional, analisando conteúdos audiovisuais a partir das perspectivas dos criadores de conteúdo, dos pares e das ferramentas de IA, com foco particular em filmes produzidos em oficinas de audiovisual supervisionadas em ambiente de pesquisa acadêmica.

Os resultados revelam divergências relevantes entre as categorizações humanas e automatizadas, destacando a influência dos algoritmos de recomendação baseados em IA na visibilidade pública e na formação de narrativas sociais. O estudo identifica vieses e limitações inerentes a essas ferramentas, abrindo margem para o reforço de estereótipos, para polarização no debate público, difusão de informações pouco confiáveis e na limitação da diversidade de perspectivas de autores e usuários por meio de critérios opacos e de motivação pouco conhecidas. Defende-se uma maior transparência e controle sobre os processos de tomada de decisão algorítmica e recomenda-se a adoção de medidas políticas para garantir ambientes digitais mais justos e inclusivos. Esta pesquisa visa contribuir para a sociologia digital, os estudos de mídia e os debates sobre políticas públicas ao propor metodologias inovadoras para melhor compreender as complexidades da formação de identidades em espaços digitais.

Palavras-chave: identidades digitais, conteúdo audiovisual, redes sociais, inteligência artificial, YouTube

DIGITAL NARRATIVES AND ARTIFICIAL INTELLIGENCE: A THREE-DIMENSIONAL COMPARATIVE ANALYSIS OF AUDIOVISUAL CONTENT

ABSTRACT: This article examines the impact of social media platforms and artificial intelligence (AI) technologies on the construction of digital identities and communication. It explores how platforms such as YouTube influence the creation, expression, and perception of identities while discussing the role of AI algorithms in shaping these interactions. The research adopts a three-dimensional comparative approach, analyzing audiovisual content from the perspectives of content creators, peers, and AI tools, with a particular focus on films produced in audiovisual workshops supervised in an academic research environment.

The findings reveal relevant divergences between human and automated categorizations, highlighting the influence of AI-based recommendation algorithms on public visibility and the formation of social narratives. The study identifies inherent biases and limitations in these tools, which can reinforce stereotypes, promote polarization in public debate, disseminate unreliable information, and limit the diversity of perspectives among authors and users through opaque and poorly understood criteria. The article advocates for greater transparency and control over algorithmic decision-making processes and recommends policy measures to ensure fairer and more inclusive digital environments. This research aims to contribute to digital sociology, media studies, and policy debates by proposing innovative methodologies to better understand the complexities of identity construction in digital spaces.

Keywords: digital identities, audiovisual content, social networks, artificial intelligence, YouTube

INTRODUÇÃO

A ascensão das plataformas digitais e das tecnologias de Inteligência Artificial (IA) transformou profundamente a comunicação e a construção de identidades no século XXI. Redes sociais como YouTube, Facebook e Instagram desempenham um papel central na disseminação de

conteúdos e na formação de discursos, impactando tanto as interações sociais quanto a forma como narrativas e identidades são criadas e percebidas, ou como as ações práticas da vida social são transformadas como já antecipava McLuhan (2019). A popularização das ferramentas de IA adiciona complexidade a esse cenário, ao influenciar diretamente quais conteúdos são promovidos, suprimidos ou ressignificados, moldando o espaço discursivo nas redes sociais.

O impacto das redes sociais na formação de identidades tem sido amplamente discutido por teóricos como Naím (2007), que analisou o efeito democratizador do YouTube, e Pariser (2011), que alertou para os riscos das "bolhas de filtro" criadas por algoritmos. Outros estudiosos, como Dubar (2005), Goffman (1988) e Castells (2000, 2009), investigaram a construção de identidades em contextos sociais, enquanto Hall (1996) e Giddens (1991) destacaram a natureza fragmentada e reflexiva das identidades na modernidade tardia. Essas abordagens ajudam a entender como as identidades são moldadas por interações sociais e fatores externos, incluindo novas tecnologias.

Apesar do avanço nessas discussões, ainda há lacunas importantes na compreensão de como as ferramentas de IA influenciam a categorização e interpretação de conteúdos audiovisuais em plataformas digitais. Enquanto estudos anteriores focaram principalmente nas interações humanas, menos atenção foi dada ao papel dos algoritmos de IA na modulação dessas interações e na formação de identidades digitais de forma segmentada e personalizada. É crucial investigar como as descrições e categorizações feitas por autores e seus pares se comparam às análises automatizadas realizadas por IA em plataformas como o YouTube, especialmente quanto às suas implicações para usuários, autores e o público em geral, no que tange a políticas públicas e regulamentações para melhor controle e transparência no uso de dados pessoais.

Neste contexto, a presente pesquisa busca preencher essa lacuna ao realizar uma análise comparativa de conteúdos audiovisuais, considerando as perspectivas dos autores, dos pares e das ferramentas de IA. Foram analisados 125 filmes produzidos em um laboratório acadêmico, com foco em 30 filmes amadores de 2024, cujas análises e categorizações foram feitas tanto por humanos quanto por algoritmos de IA do YouTube Analytics. O estudo explora as convergências e divergências nas interpretações, destacando os vieses e limitações das ferramentas de IA, e oferece insights sobre a necessidade de políticas públicas que promovam maior transparência e controle sobre os algoritmos.

Dessa forma, este estudo diferencia-se ao integrar uma abordagem comparativa entre análises humanas e automatizadas, uma perspectiva relativamente nova na sociologia digital e nos estudos de comunicação. Suas contribuições buscam aprofundar a compreensão das dinâmicas de construção de identidades em ambientes digitais e propor estratégias para uma regulação mais eficaz e justa do uso de IA em plataformas sociais.

1 REFERENCIAL TEÓRICO

1.1. O PAPEL DAS REDES SOCIAIS NA COMUNICAÇÃO E NA FORMAÇÃO DE IDENTIDADES

O papel das redes sociais na revolução dos processos de comunicação e na formação de identidades foi inicialmente apontado por Moisés Naím (2007) em "The YouTube Effect", um dos primeiros artigos sobre o tema a ganhar projeção. Naím discute como plataformas como o YouTube democratizaram a produção e disseminação de conteúdos, permitindo que indivíduos comuns influenciassem narrativas políticas e sociais. Ele compara esse fenômeno ao "efeito CNN" dos anos 1990, quando a cobertura em tempo real da Guerra do Golfo trouxe a realidade dos conflitos diretamente para o público, mas destaca que o YouTube vai além ao permitir que qualquer pessoa com uma câmera crie e compartilhe conteúdo quase instantaneamente, rompendo as barreiras impostas pelos meios de comunicação tradicionais. A previsão de Naím (2007) se concretizou ao longo das décadas seguintes, com as redes sociais desempenhando papéis centrais em eventos políticos e sociais significativos, como revoluções e protestos transmitidos em tempo real. Entre a publicação do autor até os dias hoje passamos por situações diversas como primaveras árabes, revoluções coloridas e golpes de estado, quase sempre ancorados em mobilizações nas redes sociais. Seja por sua capacidade de disseminação de informação e de desinformação no contexto de crises políticas (Reuter et al., 2020), as redes sociais têm se mostrado ferramentas essenciais para facilitar e planejar mobilizações (Hamanaka, 2020), especialmente pela sua abrangência e rapidez, sem a necessidade de uma liderança centralizada (Tarafdar; Ray, 2021), e tal como pode-se observar no atual contexto político global estão sujeitas a interferências de seus gestores como observado no caso da proibição decretada pelo Supremo Tribunal Federal do uso da Plataforma X (antigo Twitter) por usuários brasileiros devido à recusa em definir um responsável legal no território brasileiro, bem como de cumprir com decisões jurídicas do país (Santos, 2024).

Pariser (2011) analisa como as redes sociais personalizam o conteúdo para os usuários, criando bolhas informacionais que limitam a diversidade de perspectivas a que esses usuários são expostos e aponta para a necessidade de as plataformas digitais contemporâneas serem mais transparentes em seus processos de tomada de decisão. Pariser sugeriu há mais de uma década que essas empresas adotem práticas que permitam ao público compreender como suas preferências e informações estão sendo utilizadas para moldar a experiência online, destacando a importância da transparência e responsabilidade para manter a confiança pública nas plataformas de mídia (Pariser, 2011).

As redes sociais, conforme discutido por Naím (2007) e Pariser (2011), possibilitaram uma maior participação do público na criação e disseminação de conteúdo, mas também abriram espaço para a propagação de desinformação. Burgess e Green (2018), destacam essa dualidade em seu estudo sobre o YouTube, observando como a plataforma por um lado facilita a produção de conteúdos bem fundamentados, mas por outro, enfrenta desafios significativos na moderação de conteúdos potencialmente prejudiciais.

1.2. A CONSTRUÇÃO DE IDENTIDADES EM PLATAFORMAS DIGITAIS

A construção de identidades em plataformas digitais como o YouTube é um processo dinâmico que envolve tanto a auto-orientação quanto a atribuição de identidade pelos outros e reflete a complexidade e a interculturalidade da atuação nas plataformas digitais contemporâneas. O YouTube, com seu acesso global e suas características de (pseudo-)anonimato e (pseudo-)familiaridade, oferece um ambiente rico para a interculturalidade, onde identidades são continuamente negociadas e expressas através de recursos linguísticos nas interações via comentários, ao mesmo tempo em que permitem diferentes níveis de civilidade e incivilidade (Yun et al., 2020). Um aspecto facilitador desse processo é a possibilidade de trocar de código (code-switching), ou seja, de se utilizar diferentes idiomas nos comentários de publicações online, o que permite que os usuários “co-constuam” e negociem identidades culturais e que experimentem um sentimento de pertença a uma comunidade, mesmo que virtual (Wentker; Schneider, 2022).

Além das identidades virtuais que vão construindo nas comunidades virtuais em que participam, os indivíduos podem incorporar as experiências dessas plataformas em suas identidades fora do ambiente online, processo esse que se conecta à exploração da própria identidade durante a

adolescência e o início da vida adulta, influenciando quem eles são e quem desejam se tornar (Soh et al., 2024, Wentker; Schneider, 2022).

A comunicação mediada por computador (CMC - computer-mediated communication) desempenhou um papel central em diversas mobilizações sociais desde o início dos anos 2010. A revisão de literatura conduzida por Priante et al. (2018) examina estudos que consideram duas perspectivas distintas sobre o processo de mobilização: de um lado, a construção de uma identidade coletiva, vista como um elemento essencial para mobilizações que emergem das interações nas plataformas online; de outro, a mera existência de conectividade entre os participantes das comunidades virtuais.

Como ferramenta de construção de identidade a comunicação online facilita a identificação com pautas, ideias, linguagem e modos de ação, o que permite a concretização de mobilizações em torno de uma variedade de causas — políticas, ideológicas, religiosas e culturais. O espaço e os recursos oferecidos pelas redes sociais para a expressão, negociação e construção de identidades sociais possibilitam a formação de identidades coletivas que impulsionam ações, campanhas e protestos de maneira ágil e eficiente (Priante et al., 2018).

Estudo realizado por Spann et al. (2023), que analisou a campanha online do movimento #stopthesteal, que vigorou durante a eleição presidencial dos EUA em 2020, os resultados mostram que a rede se tornou mais coesa e a similaridade de tópicos aumentou dentro das comunidades antes e logo após as eleições e o ataque ao Capitólio dos EUA e funcionou um como catalisador para o evento (Spann et al., 2023).

A influência dos algoritmos sobre as interações e a visibilidade dos conteúdos gera preocupações sobre a transparência e o controle que os criadores de conteúdo têm sobre suas próprias narrativas e identidades e sobre quais são os seus públicos. A construção de identidade online é fortemente influenciada por fatores como a indexicidade (ex. a menção explícita de categorias e rótulos atribuídos) e as oportunidades de interação que permitem que as pessoas se definam em relação aos outros (Bucholtz; Hall, 2005).

Peralta et al. (2021) apresenta um quadro teórico geral que conecta sistematicamente modelos de dinâmica de opinião, estrutura de redes sociais e filtragem de conteúdo. Todos os modelos

contemplados em seu estudo indicam um regime de polarização de opiniões impulsionado pelo viés algorítmico e pela estrutura modular da rede. No entanto, o papel da filtragem de conteúdo é complexo: nas interações de pares, leva à polarização, enquanto nas interações em grupo promove a coexistência de opiniões (Peralta et al, 2021).

A seguir apresentamos uma análise de filmes amadores produzidos em laboratório de audiovisual, de modo a comparar como esses materiais, publicados no YouTube, foram percebidos e categorizados em três dimensões: de um lado, i) seus próprios autores, ii) os autores de outros filmes realizados nas mesmas condições e período, que são chamados de pares, e de outro iii) os termos extraídos da plataforma Analytics do YouTube de modo indireto para viabilizar uma análise de correspondência entre os termos mobilizados nestas três dimensões. Utilizamos essas categorizações para confrontar os 30 filmes de produção mais recente — criados e analisados por 106 pessoas — com as temáticas de filmes sugeridas pelo algoritmo do YouTube. Especificamente, investigamos quais temas de filmes de terceiros presentes na plataforma são recomendados ao espectador como links para os filmes da playlist da oficina, esta relação artificial que a plataforma estabelece ao sugerir filmes ao final de outras exibições é mediada por parâmetros e algoritmos (Google, 2024) que são definidores e categorizadores dos materiais que são relacionados com base em critérios pouco conhecido e pouco transparentes.

1.3. LEGISLAÇÕES SOBRE OS USUÁRIOS E AUTORES DE PLATAFORMAS DIGITAIS E DAS REDES SOCIAIS

De acordo com a Lei Geral de Proteção de Dados Pessoais (LGPD) – Lei nº 13.709/2018 (Brasil, 2018), os usuários têm o direito de solicitar ao YouTube informações sobre como seus dados pessoais são processados, incluindo o uso de algoritmos para categorizar perfis ou conteúdos, como filmes. A LGPD assegura o Direito de Acesso (Art. 18, II), permitindo que os usuários obtenham informações detalhadas sobre o tratamento de seus dados, incluindo qualquer categorização automatizada.

Além disso, o Direito à Explicação sobre Decisões Automatizadas (Art. 20) concede aos usuários o direito de solicitar revisão de decisões tomadas exclusivamente por algoritmos que afetem seus interesses, como a categorização de perfis ou conteúdos (Brasil, 2018). O princípio da

transparência (Art. 6, VI, e Art. 9) exige que as empresas, como o YouTube, forneçam informações claras sobre o uso de algoritmos que categorizam dados pessoais (Brasil, 2018).

Os usuários também têm o Direito de Informação (Art. 18, VII) para solicitar detalhes sobre com quem seus dados são compartilhados e como são processados (Brasil, 2018). Para exercer esses direitos, é possível enviar uma solicitação formal ao YouTube, pedindo esclarecimentos sobre a categorização automatizada de dados e seus critérios. Caso uma decisão automatizada afete significativamente o usuário, ele pode solicitar uma revisão humana, conforme previsto pela LGPD (Brasil, 2018).

Apesar desses direitos, a aplicação prática nas plataformas digitais enfrenta desafios, pois as empresas geralmente fornecem respostas gerais, limitadas pela proteção de segredos comerciais. A transparência ainda depende do desenvolvimento regulatório e da atuação de entidades como a Autoridade Nacional de Proteção de Dados (ANPD) no Brasil (ANPD, 2024).

2 CONTEXTO DA PESQUISA

Entre 2016 e 2019, foram realizadas oficinas de audiovisual presenciais no Laboratório de Pesquisa Social da Universidade de São Paulo, visando integrar atividades acadêmicas com a comunidade não acadêmica interessada em produção audiovisual e questões sociais. Durante essa fase, o NUPEPA-ImaRgens organizou oito oficinas com um total de 151 participantes, cada uma com carga horária média de 54 horas.

As oficinas tinham como objetivo facilitar o acesso e a compreensão de conceitos e práticas da produção audiovisual, promovendo a troca de ideias entre professores, pesquisadores, estudantes e membros da sociedade, abordando temas acadêmicos, artísticos e sociais. Tanto os interessados em pesquisa acadêmica quanto aqueles que buscavam experiências práticas fora do ambiente acadêmico podiam se beneficiar desse espaço de aprendizado.

A partir de 2020, no contexto de investigação na área de Comunicação e Ciências Sociais no ICNOVA da Universidade Nova de Lisboa, foi desenvolvido um modelo virtual para as oficinas, aprimorando a metodologia de coleta de dados e permitindo a participação de um público geograficamente diverso. A média de participantes passou de 24 nas edições presenciais na USP para 60 nas edições virtuais, com inscritos de 17 países, incluindo Américas, África, Europa e Oriente

Médio. As oficinas virtuais utilizam formulários digitais do Google para dinâmicas individuais e de grupo, permitindo a coleta de novas camadas de dados.

Todas as oficinas são gratuitas, organizadas por pesquisadores de doutorado das universidades públicas vinculadas ao NUPEPA-ImaRgens, LAPS (USP) e ICNOVA (Universidade Nova de Lisboa). A única contrapartida exigida dos participantes é a contribuição para a pesquisa científica em andamento, por meio do preenchimento de formulários, com dados anonimizados, e a disponibilização de seus trabalhos finais nos canais das instituições envolvidas.

Os filmes produzidos pelos participantes são de caráter público e estão disponíveis nos canais do NUPEPA-ImaRgens e de seus parceiros acadêmicos. Este estudo utiliza dados do período virtual das oficinas realizadas entre 2020 e 2024, nas quais foram produzidos 95 filmes por 481 participantes. Devido ao preenchimento incompleto de alguns formulários de pesquisa, a participação efetiva variou entre 445 e 455 respondentes, dependendo do conjunto de formulários analisados e do escopo da análise. Assim, o total de participantes (N_p) pode variar conforme o bloco de questões preenchido e validado. No total, foram produzidos 127 filmes, mas apenas 125 estão publicados e disponíveis ao público ($N_f = 125$). Dos filmes analisados, 30 permitem comparações com dados do YouTube Analytics ($N_{30f} = 30$), enquanto 95 foram comparados com os termos dados por autores e pares ($N_{95f} = 95$).

3 MÉTODO

3.1 O PROCESSO DE RECOLHA DE DADOS

O processo de coleta de dados para esta pesquisa incluiu formulários digitais preenchidos individualmente e em grupo, registro de interações durante atividades e materiais produzidos. Foram utilizados dados dos questionários em duas dimensões: a) no nível de grupo, onde 3 a 5 participantes forneceram informações sobre sinopses, títulos, detalhes da produção, funções de membros e tempo despendido; b) análises individuais que cada participante fez do próprio filme e de três filmes de colegas da mesma oficina; e c) dados disponíveis na plataforma YouTube, como transcrições, dados do API e do recurso Analytics.

Os autores e pares foram convidados a descrever "sobre o que é o filme" em uma ou duas palavras. A escolha dos filmes a serem analisados pelos pares foi livre, resultando em um número variável de análises por filme, influenciado por temas de interesse, afinidades pessoais, ordem de

apresentação na playlist, entre outros fatores. As descrições feitas por autores e pares foram então comparadas para identificar relações ou divergências, considerando termos diretos, sinônimos, antônimos, ou referências indiretas de caráter causal, consequencial, contextual (sintético) ou de detalhe (analítico). Os resultados das análises independentes foram combinados em categorias mais inclusivas.

Foram comparados então os termos de cada conjunto de análises feitos por pares e autores que tinham correspondência em cada dimensão partindo da dimensão do autor, com o total de termos informados por autores e pares. O percentual observado nas tabelas de detalhe mais abaixo entre os elementos A (autores) e P (pares) corresponde ao percentual da soma dos valores correspondentes entre os dois conjuntos. Esta comparação permite extrair um percentual de correlação entre a identidade que os autores atribuem ao filme que produzem e a identidade percebida pelos seus pares. Neste estudo não foram analisados comentários ou declarações do público externo à oficina por não estar habilitado o espaço para comentários por regra de publicação do canal o que motivou a busca por métodos indiretos de cruzamento de dados a partir da base de dados do Analytics do YouTube.

A análise baseada nos dados da plataforma ocorreu de modo indireto, por meio da relação de vídeos sugeridos que direcionavam usuários para os filmes do NUPEPA-ImaRgens. Essa recomendação sugere uma correlação estabelecida por algoritmos, rotulagem e categorização de filmes, que pode ser fundamentada em análises de similaridade de conteúdo, histórico de engajamento, modelos de machine learning, feedback explícito e implícito, clusters de interesse e personalização contextual. A plataforma considera o histórico de visualizações e pesquisas, padrões de comportamento do usuário, como frequência e atualidade das visualizações, além de interações diretas, como curtidas e compartilhamentos. Utilizando aprendizado de máquina, o algoritmo identifica padrões e preferências, agrupando usuários com comportamentos semelhantes para prever interesses futuros. Assim, relaciona vídeos previamente assistidos com novas sugestões, de acordo com categorias e temas de potencial interesse, sem necessariamente detalhar os critérios específicos dessa categorização (Arthurs et al., 2018; Peralta et al, 2021; Google, 2024; ImaRgens, 2024).

O passo seguinte foi obter do recurso Analytics do YouTube os dados sobre a quantidade e a origem das visualizações que os filmes da playlist da oficina obtiveram.

No painel de análise do YouTube não é possível identificar quais filmes são sugeridos a partir de algum dos filmes do NUPEPA-ImaRgens, também não é possível identificar categorias temáticas ou de qualquer tipo que indique de que modo o YouTube interpreta e categoriza o material, nem a indicação de parâmetros, métricas e outros dados que possam indicar o motor de recomendação de conteúdos. Para que uma playlist possa ter recursos do Analytics habilitados é necessário que haja uma quantidade mínima de visualizações. Devido a esta limitação somente foi possível comparar os dados do painel de análise do YouTube referente aos 30 filmes resultantes da última oficina concluída em julho de 2024 que teve um total visualizações de 10.181 no período de 1 de julho até 31 de agosto de 2024.

Os filmes analisados neste estudo são de acesso público. Embora os dados dos participantes sejam anônimos, os autores podem ser identificados pelos créditos dos filmes que produziram, mas como suas identidades individuais são tratadas de forma agregada, isso assegura o anonimato quanto aos dados pessoais. As informações fornecidas conscientemente pelos participantes, através de questionários, diferem daquelas coletadas pelas plataformas digitais, que geram categorizações de forma opaca, sem permitir que os usuários conheçam como seus dados são processados. Os participantes podem gerar automaticamente uma versão de suas respostas para controle pessoal e solicitar informações ao NUPEPA-ImaRgens a qualquer momento, algo que não é facilmente replicável nas plataformas digitais. Além disso, como os dados coletados são utilizados para pesquisa científica, os respondentes têm a oportunidade de acessar os resultados das investigações, sempre apresentados de forma anonimizada e agregada, em estudos e publicações derivados desta base de dados.

3.2 REPRODUTIBILIDADE DO EXPERIMENTO

Esta análise pode ser replicada seguindo os seguintes procedimentos: i) acompanhamento, produção ou curadoria coletiva de filmes amadores, preferencialmente realizados por grupos cujos integrantes não se conheciam anteriormente; ii) aplicação de questionário para coletar características sociodemográficas e outros indicadores de interesse, como preferências temáticas; iii) promoção de uma mostra dos filmes produzidos para os autores e pares do mesmo curso ou programa; iv) aplicação de questionário específico para obter a percepção dos autores e pares sobre diferentes dimensões analíticas, usando uma escala Likert de 1 a 5 para avaliar o quanto consideram que o filme explora aspectos lógicos (explicações causais e objetivas), emocionais (apelos emocionais e subjetivos), e

morais/éticos (relativos às normas sociais), além de solicitar um resumo com a pergunta: "Sobre o que é o filme?" e uma definição em 1 ou 2 palavras — esta última sendo a principal base para análise dos pares; v) comparar os termos usados por autores e pares com as descrições e categorizações geradas automaticamente pelas plataformas ou buscar marcadores indiretos, como as sugestões de vídeos relacionados pelo YouTube entre materiais de terceiros e os filmes considerados no experimento. Os itens i, iv, v e vi são os mínimos necessários para reproduzir esta pesquisa.

4 ANÁLISE DOS DADOS

Após a organização e padronização dos dados, foram realizadas análises detalhadas dos termos utilizados por autores, pares e sugestões do YouTube. Com base na amostra total das oficinas de 2020 a 2024 ($N_f = 125$), foi possível gerar correlações entre as dimensões de Autores (A) e Pares (P). Para os filmes produzidos em 2024 ($N_{30f} = 30$), que atingiram o mínimo de visualizações exigido pelo YouTube, foi possível observar também a correlação com as Sugestões do YouTube. A Tabela 1 apresenta o nível de correlação entre termos das três dimensões analisadas: AP (Autores e Pares), AS (Autores e Sugestões do YouTube), PS (Pares e Sugestões do YouTube), e APS (Autores, Pares e Sugestões do YouTube). As correlações foram baseadas na interseção dos termos, priorizando, sempre que possível, a dimensão das Sugestões do YouTube, seguida pela dos Autores e, por fim, pelos Pares.

Amostra	ID Filme	Filme	% Relac AP	% Relac AS	% Relac PS	% Relac PAS
Nf=30	1110201	ANINHO	65.79%	33.33%	70.00%	7.14%
Nf=30	1110202	SONHOS KANÃ MIHAY	80.00%	40.00%	62.86%	12.50%
Nf=30	1110203	AMBIENTALISMO EM PORTO ALEGRE	91.67%	0.00%	0.00%	0.00%
Nf=30	1110204	TELA DELAS	88.89%	40.00%	50.00%	17.39%
Nf=30	1110205	AS ABELHAS VOARAM	53.33%	0.00%	0.00%	0.00%
Nf=30	1110206	DESTINOS	80.00%	16.67%	0.00%	20.00%
Nf=30	1110207	A HISTÓRIA DA ÁGUA	100.00%	41.67%	58.33%	26.32%
Nf=30	1110208	ESCOLA CÍVICO MILITAR: A QUEM INTERESSA?	92.86%	0.00%	0.00%	0.00%
Nf=30	1110209	MULHERES CAMPONESAS EM DEFESA DA TERRA E DAS PLANTAS MEDICINAIS	92.59%	37.50%	58.62%	15.63%
Nf=30	1110210	POR TRÁS DAS FITAS	100.00%	44.44%	73.68%	17.39%
Nf=30	1110211	OPARÁ - O VELHO CHICO: ÁGUAS, MARGENS E GENTES	69.57%	40.00%	69.57%	17.86%
Nf=30	1110212	CASA DE SONHOS	94.12%	44.44%	75.00%	19.05%
Nf=30	1110213	BOLERO E LAMBADA: NO CORAÇÃO DO PARÁ	100.00%	0.00%	0.00%	0.00%
Nf=30	1110214	FRONTEIRA	47.06%	0.00%	0.00%	0.00%
Nf=30	1110215	RIO GRANDE ENORME	90.00%	25.00%	50.00%	26.67%
Nf=30	1110216	ANÉIS DE COCO: ONDE A ENCHENTE E A SOLIDARIEDADE SE ENCONTRAM	100.00%	0.00%	0.00%	0.00%
Nf=30	1110217	A CIDADE E O MOVIMENTO DA ÁGUA	81.82%	0.00%	0.00%	0.00%
Nf=30	1110218	TAMBORES DE RESISTÊNCIA	100.00%	37.50%	72.22%	23.81%
Nf=30	1110219	MAIS UMA VEZ ACORDO	76.47%	0.00%	0.00%	0.00%
Nf=30	1110220	LITORAL DO PARANÁ: TERRITÓRIO MÚLTIPLO, DIVERSO E ANCESTRAL	0.00%	0.00%	0.00%	0.00%
Nf=30	1110221	ÁGUA SEM CASA CIDADE SEM ÁGUA	66.67%	28.57%	0.00%	62.50%
Nf=30	1110222	O QUE FICA DEPOIS DA ENCHENTE?	91.67%	42.86%	22.22%	14.29%
Nf=30	1110223	RUÍNAS	50.00%	0.00%	0.00%	0.00%
Nf=30	1110224	RUÍDOS DO POEMA	37.50%	12.50%	20.00%	15.38%
Nf=30	1110225	CAFÉ EM PÓ-É-CHICAS	100.00%	33.33%	50.00%	6.25%
Nf=30	1110226	INTERVENÇÕES URBANAS	100.00%	44.44%	37.50%	33.33%
Nf=30	1110227	ENTRE NÓS E NOSSAS COISAS	80.00%	0.00%	0.00%	0.00%
Nf=30	1110228	ANGÚSTIA	90.00%	44.44%	27.27%	6.67%
Nf=30	1110229	COTIDIANO ENCANTADO	78.57%	12.50%	12.50%	5.26%
Nf=30	1110230	SENTIDOS CRUZADOS	60.00%	20.00%	20.00%	13.33%

Tabela 1. Relação de filmes analisados nas três dimensões: pares, autores e sugestões do YouTube

As células marcadas em vermelho na tabela 1 acima apresentam alto nível de correlação entre as dimensões analisadas, as marcadas em azul apresentam baixa correlação. Podemos notar que os termos dados por autores e pares proporcionam alto nível de correlação, enquanto os obtidos pelo Analytics do YouTube com base nas sugestões de vídeos apresentam baixa correlação dos termos.

Os dados de base destas análises têm como base os registros disponíveis no repositório do GitHub e indicam o quão significativa, do ponto de vista estatístico, é a amostra utilizada como dados de base para testar a correlação entre os termos das dimensões estudadas¹.

¹ As colunas "Relac P-S (restrito)" e "Relac A-S em A (restrito)" apresentaram um valor de p (Correlação de Spearman) de 0,0400, indicando uma associação significativa entre essas variáveis, sugerindo uma relação consistente entre as colunas de convergentes de Pares e Sugestões do YouTube (P-S) e entre Autores e Sugestões da plataforma (A-S) em ambos os contextos. A correlação entre "Relac P-S (restrito)" e "Relac PAS (restrito)" mostrou um valor de p de 0,0001, evidenciando uma forte correspondência estatística entre os termos dados pelos Pares e Sugestões da Plataforma (P-A) e a intersecção dos termos observados nas três dimensões. Outra correlação significativa foi encontrada entre "Relac A-S em A (restrito)" e "Relac PAS (restrito)", com um valor de p de 0,0200, sugerindo que uma convergência entre os termos dados por Autores e Pares, bem como na intersecção destas dimensões. Notavelmente, a associação entre "Relac A-S em A (restrito)" e "Relac A-S em A (amplo)" teve um valor de p de 0,0001, indicando uma correlação muito forte, como esperado, dado que ambas

Na tabela 2 a seguir é possível observar os termos convergentes dados pelos pares, pelos autores e pelos temas observados nos filmes sugeridos pelo YouTube em relação ao filme que teve maior número de visualizações no YouTube. A tabela permite observar as quantidades totais de Termos dadas pelos Pares (P), pelos Autores (A) ou obtidos por meio das Sugestões do YouTube (S), sendo que para esta última dimensão foram consideradas as cinco primeiras sugestões dadas pelo YouTube em ordem decrescente de visualizações geradas para o filme de destino. Não é possível observar nesta tabela os totais de relações parciais entre autores, pares e sugestões do YouTube.²

Exemplo de análise do filme 1110215						
QTD P.	ANÁLISE DOS PARES	Relacionados	A-P	P-S		
1	Calamidade e perseverança	1	1	1		
1	Devastação e Sobrevivência	1	1	1		
1	Enchente	1	1	1		
1	Memória e superação.	1				
1	Resiliência e Solidariedade.	1	1			
1	Solidariedade	1	1			
1	um recorte narrativo-exploratório-realista por meio de coleta de depoin	1	1			
QTD A.	ANÁLISE DOS AUTORES	Relacionados	A-P	A-S		
1	cooperação	1	1		1	
1	enchentes e trabalho voluntário	1	1		1	1
1	Realista	1	1			
QTD. S.	TEMAS SUG	Relacionados	P-S	A-S	PAS	
1	violência, mortes por crime, psicopata					
1	desastres naturais, intrigas governamentais, ficção científica	1		1	1	1
1	desastres naturais, fim do mundo, crise climática	1		1	1	1
1	desastres naturais, fim do mundo, ficção científica	1		1	1	1
1	desigualdades sociais, busca pela felicidade, conflitos familiares	1		1	1	1
15	27% de correspondência entre as 3 dimensões de análise (15/4)	4				

Tabela 2. Exemplo de análise comparativa manual entre os termos por dimensão P, A ou S

A tabela abaixo mostra um recorte com 1 dos 125 registros agregados para as dimensões analíticas deste estudo. O registro em destaque corresponde ao da tabela anterior referente ao filme 1110215^{3 4}. Este registro de exemplo permite comparar os dados de correlação detalhados na tabela

as medidas envolvem dimensões restritas de articulação social. Esses resultados demonstram uma consistência significativa entre as variáveis restritas analisadas, sugerindo uma robusta validade interna das relações no conjunto de dados. As correlações significativas reforçam a confiabilidade dos dados, especialmente considerando a natureza específica e contextualizada das relações sociais em ambientes limitados. Detalhes adicionais estão disponíveis no GitHub da publicação: <https://github.com/nupepa/48ANPOCS>.

² Os detalhes dos cálculos dos filmes podem ser conferidos nas tabelas do repositório do GitHub: <https://github.com/nupepa/48ANPOCS>

³ Endereço de acesso às bases de dados deste estudo: <https://github.com/nupepa/48ANPOCS>

⁴ Os dados estão disponíveis em detalhe no GitHub tanto na tabela com todos os registros detalhados (* T-Anon.DadosParticipTermos127f do arquivo de Tabelas de Detalhamento e Apoio) quanto na tabela de Exemplos do arquivo (* Dados Descritivos e Amostra Utilizada).

2 com os demais índices e marcadores utilizados para estabelecer a proporção de correlação entre os termos.

Amostra	ID Filme	Filme	Qtd	QTD_P	QTD_A	QTD_S	TOTAL PAS	Relac A-P em A	Relac P-A em P	Relac P-A (restrito)	Total PA	Relac P-S (restrito)	Relac P-S em P (amplo)	Relac A-S em A (restrito)	Relac A-S em A (amplo)	Relac PAS (restrito)	% Relac AP	% Relac AS	% Relac PS	% Relac PAS
Nf=30	1110215	RIOGRANDE ENORME	1	7	3	5	15	3	6	9	10	6	6	2	2	4	90.00%	25.00%	50.00%	26.67%

Tabela 3. Cálculos utilizados para estabelecer a correlação entre as dimensões P, A e S do filme 1110215.

A seguir apresentamos dados fornecidos pelo YouTube sobre o filme com maior número de visualizações da playlist de filmes produzidos na oficina realizada pelo NUPEPA-ImaRgens de março a julho de 2024, de 6 de julho a 31 de agosto de 2024. O filme Rio Grande Enorme foi publicado no dia 6 de julho de 2024 na playlist no canal do NUPEPA-ImaRgens contendo os filmes produzidos pela oficina de audiovisual realizada. Os dados apresentados foram extraídos de um dos painéis do recurso Analytics do YouTube que informa a origem das visualizações, e uma destas origens são filmes anteriormente vistos por espectadores que sugerem ao espectador um novo filme. Destacam-se os cinco filmes do YouTube que mais vezes sugeriram este filme da oficina, cada um correspondendo a uma cor no Gráfico X. São eles: o longa-metragem de ficção “A morte pede carona – Filme completo e dublado”, o longa-metragem de ficção científica “O dia em que a Terra parou”, um vídeo amador com uma seleção de filmes sobre desastres naturais com título “TOP 35 minutes of natural disasters”, “A Onda” (filme norueguês de 2015 sobre um tsunami) e o longa-metragem ficcional “O PÁSSARO AZUL”.

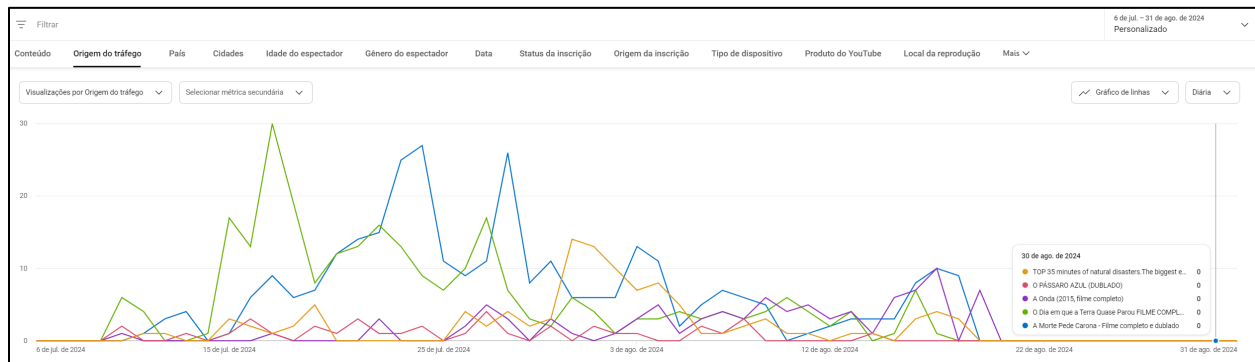


Gráfico 1. Visualizações diárias do filme 1110215 por origem de tráfego sugerido pelo YouTube⁵.

⁵ Visualizações entre o período de 06 de julho a 31 de agosto.

O total de visualizações do filme 1110215, que tem o título “Rio Grande Enorme”, originadas das cinco principais sugestões de filmes de terceiros do YouTube pode ser visto na tabela 4 a seguir.

Origem do tráfego > Vídeos sugeridos	Visualizações ↓	Tempo de exibição (horas)	Duração média da visualização	Impressões	Taxa de cliques de impressões
<input type="checkbox"/> Total	2.521	146,9	3:29	109.632	2,1%
<input type="checkbox"/> A Morte Pede Carona - Filme completo e dublado	312 12,4%	15,5 10,5%	2:58	22.184	1,3%
<input type="checkbox"/> O Dia em que a Terra Quase Parou FILME COMPLETO DUBLADO Fil...	267 10,6%	14,5 9,9%	3:15	17.484	1,4%
<input type="checkbox"/> TOP 35 minutes of natural disasters.The biggest events in world. T...	108 4,3%	4,0 2,7%	2:12	2.562	3,9%
<input type="checkbox"/> A Onda (2015, filme completo)	92 3,7%	4,4 3,0%	2:53	5.605	1,6%
<input type="checkbox"/> O PÁSSARO AZUL (DUBLADO)	37 1,5%	2,1 1,4%	3:21	3.716	0,9%

Tabela 4. Número de visualizações originadas da sugestão de YouTube por filme.

A Tabela 4 mostra que quatro dos cinco principais geradores de visualizações para o documentário Rio Grande Enorme são filmes de ficção. Dois desses filmes, O Dia em que a Terra Quase Parou e A Onda, abordam temas relacionados a desastres naturais, correlatos ao tema do documentário. Os outros dois, A Morte Pede Carona (sobre um psicopata) e O Pássaro Azul (sobre a jornada de uma criança em busca da felicidade), responsáveis por 349 visualizações (13,9% do total de sugestões do YouTube), não possuem relação temática com o filme da oficina, mas ainda assim o sugeriram aos espectadores. O vídeo amador TOP 35 Minutes of Natural Disasters, publicado em um canal sensacionalista, gerou 108 visualizações e teve a maior taxa de conversão de cliques por impressão (3,9%), como indicado na última coluna da Tabela 4. Conteúdos sensacionalistas, como esse, geralmente geram alto engajamento e são associados à polarização e à desinformação (Bryant, 2020).

Embora não seja possível determinar os algoritmos que o YouTube utilizou para gerar essas associações, os dados sugerem vieses de popularidade, como no caso dos filmes de ficção com temas distintos, e de homogeneidade, como nos filmes sobre desastres naturais (Nikolov et al., 2019). Além disso, o vídeo pode ter sido mais sugerido durante o período das enchentes no Rio Grande do Sul, quando foi publicado.

O YouTube utiliza uma combinação de dados históricos, padrões de comportamento e feedbacks explícitos/implícitos (como curtidas e compartilhamentos) para criar marcadores que categorizam o usuário, atribuindo-lhe uma identidade sem que este possa controlá-la ou gerenciá-la. A plataforma armazena e usa essas identidades para seus algoritmos, sem informar o usuário se ele é considerado alguém que “gosta de culinária”, é “afeito à violência” ou “gosta de conteúdo erótico”.

De forma semelhante, autores que submetem seus materiais a canais como o NUPEPA-ImaRgens também não têm clareza sobre como a identidade de seus conteúdos é gerida pela plataforma. O aprendizado de máquina utilizado para definir categorizações não permite que os autores conheçam os parâmetros usados para sugerir conteúdos ou publicidade.

Esta inconsciência geral do público e dos autores sobre como as suas próprias identidades são definidas e geridas pela plataforma, e que pode ser extrapolada para qualquer ambiente de Inteligência Artificial que utiliza aprendizado de máquina para relacionar-se com usuários voluntários ou involuntários da plataforma, impede que tenham controle ou agência sobre suas próprias identidades. É o que Cunningham et al. (2016) chama de “o misterioso”, uma vez que os algoritmos estão em constante atualização, os produtores de vídeos do YouTube vão sempre se deparar com resultados imprevisíveis e pouco controláveis. Se as legislações nacionais, como a Constituição Federal brasileira, protegem os cidadãos quanto à sua imagem e privacidade, de outro lado há o risco que de em escala muito maior a “meta-imagem” dos cidadãos no universo virtual esteja tão exposta quanto no período medieval anterior ao pleno estado de direito.

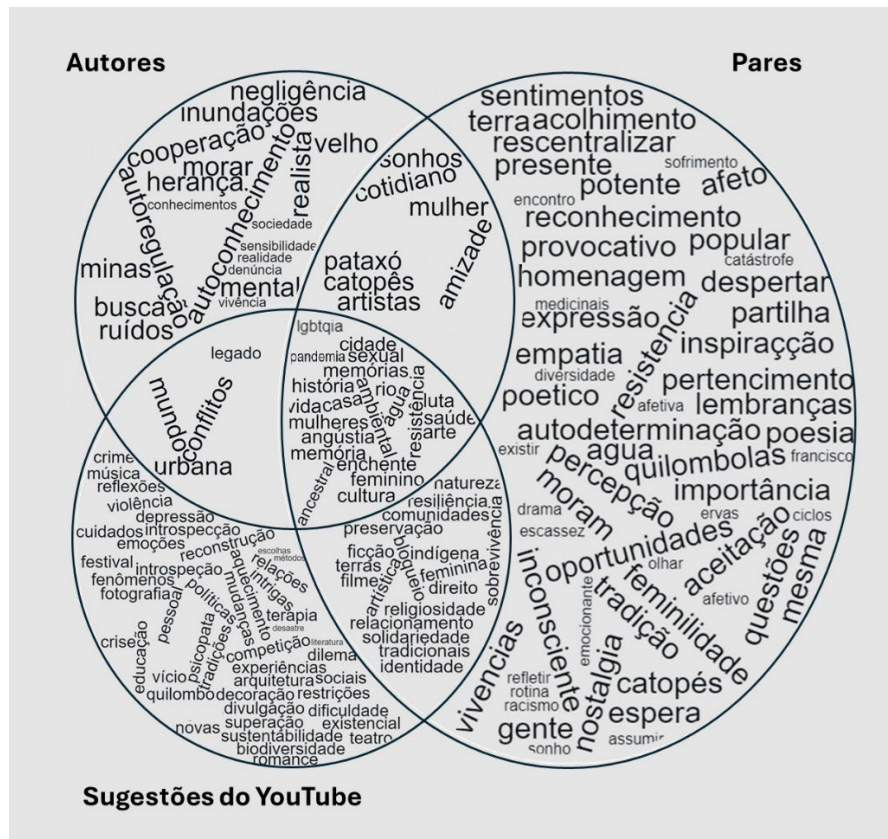


Gráfico 2. Intersecção e relação de termos padronizados por AI para dimensões P, A e S

A imagem apresentada mostra uma categorização de termos realizada com o auxílio de Inteligência Artificial (IA), destacando a interseccionalidade entre as principais temáticas atribuídas aos filmes analisados por três grupos: Autores, Pares e Sugestões do YouTube. Cada círculo representa as palavras mais associadas aos filmes de cada grupo, com as áreas de interseção indicando convergências temáticas. Essa análise, embora imperfeita, oferece uma visão elucidativa sobre como diferentes dimensões analíticas informadas pelos participantes humanos convergem ou divergem dos resultados oferecidos pela IA para definir termos e categorias, revelando tanto semelhanças quanto diferenças na relação entre conteúdos. No círculo dos termos padronizados a partir dos dados de filmes sugeridos pelo YouTube é possível ver termos como violência, psicopatia, emoções, que sugerem um conteúdo com maior apelo emocional do que aquele percebido por autores e pares. Apesar da simplificação inerente ao processo automatizado de categorização, que pode generalizar nuances captadas por especialistas humanos, o mapeamento evidencia as potencialidades e limitações do uso de IA na organização e interpretação de termos.

CONSIDERAÇÕES FINAIS

Este estudo explorou as diferenças e semelhanças nas descrições e categorizações de conteúdos audiovisuais feitas por autores e seus pares (humanos) e ferramentas de Inteligência Artificial (IA), tal qual os algoritmos de recomendação do YouTube. Os resultados mostram que, embora haja convergências, as interpretações automatizadas frequentemente divergem das humanas, especialmente quando algoritmos priorizam conteúdos de alto engajamento, como temas sensacionalistas ou emocionalmente carregados. Essas discrepâncias indicam que os algoritmos de IA influenciam ativamente a visibilidade e percepção pública dos conteúdos, reforçando estereótipos e limitando a diversidade de perspectivas.

Além disso, o estudo identificou vieses e limitações significativas nas ferramentas de IA utilizadas para a categorização e promoção de conteúdos digitais. Tais vieses têm implicações profundas na construção de identidades e narrativas em plataformas digitais, moldando o acesso e a visibilidade das informações, contribuindo para a criação de "bolhas de filtro" e para a polarização social. Este cenário ressalta a necessidade urgente de políticas públicas e regulamentações que promovam maior transparência e controle sobre os algoritmos que governam essas plataformas.

Os resultados também destacam que, apesar das contribuições deste estudo para compreender o impacto dos algoritmos de recomendação na construção de identidades digitais, permanecem lacunas significativas que devem ser exploradas em futuras pesquisas. É importante examinar pela perspectiva das identidades sociais, coletivas e individuais como diferentes grupos demográficos percebem e são afetados por essas práticas de categorização automatizada, considerando variáveis como idade, gênero, localização geográfica e perfil socioeconômico.

Adicionalmente, pesquisas futuras poderão investigar de forma mais dinâmica a interação entre usuários e algoritmos de recomendação, explorando como as identidades dos usuários moldam e são moldadas por essas ferramentas em tempo real. O desenvolvimento de métodos de auditoria mais transparentes e acessíveis para compreender melhor os impactos sociais e culturais das práticas de categorização automatizada também é uma área promissora para pesquisa.

Os resultados sugerem que algoritmos de recomendação do YouTube, influenciam identidades e narrativas digitais, moldando o espaço discursivo de forma complexa e frequentemente opaca. Compreender essas influências é essencial para criar políticas justas e eficazes que garantam transparência e equidade no uso de tecnologias de IA em plataformas digitais. Estudos como este contribuem para debates na sociologia digital, comunicação e políticas públicas, propondo uma abordagem comparativa original entre análises humanas e automatizadas e destacando a necessidade de uma governança mais responsável das ferramentas de IA.

É fundamental promover discussões públicas, pesquisas e legislações que assegurem o controle de autores e usuários sobre seus dados, especialmente aqueles relacionados às suas identidades. Leis como a Lei Geral de Proteção de Dados Pessoais (LGPD) no Brasil e o Regulamento Geral sobre a Proteção de Dados (GDPR) na União Europeia garantem direitos importantes aos usuários, mas ainda enfrentam desafios para assegurar que as plataformas sejam transparentes sobre como categorizam e rotulam identidades e conteúdos (European Parliament and Council, 2016, 2022). As plataformas influenciam diretamente a imagem dos usuários com base em critérios pouco claros, limitando a circulação de conteúdos e expondo-os a discursos de ódio. Portanto, é essencial defender políticas que promovam transparência e controle sobre os parâmetros usados pelas plataformas para categorizar e rotular conteúdos, garantindo um ambiente digital inclusivo e democrático, onde a diversidade de narrativas possa ser expressa de maneira justa e transparente.

REFERÊNCIAS

Arthurs, J., Drakopoulou, S., & Gandini, A. (2018). Researching YouTube. *Convergence*, 24(1), 3-15. <https://doi.org/10.1177/1354856517737222>

Autoridade Nacional de Proteção de Dados (ANPD). (2024). Página inicial. Autoridade Nacional de Proteção de Dados. <https://www.gov.br/anpd/pt-br>

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4-5), 585–614. <https://doi.org/10.1177/1461445605054407>

Brasil. (2018). Lei Geral de Proteção de Dados Pessoais, Lei n.º 13.709/2018. http://www.planalto.gov.br/ccivil_03/ato2015-2018/2018/lei/113709.htm

Brasil. (2014). Marco Civil da Internet, Lei n.º 12.965/2014. http://www.planalto.gov.br/ccivil_03/ato2011-2014/2014/lei/112965.htm

McLuhan, M. (2019). The medium is the message (1964). In *Crime and media* (pp. 20-31). Routledge

Bryant, L. V. (2020). The YouTube algorithm and the alt-right filter bubble. *Open Information Science*, 4(1), 85–90. <https://doi.org/10.1515/opis-2020-0007>

Burgess, J., & Green, J. (2018). *YouTube: Online video and participatory culture* (2nd ed.). Polity Press.

California Consumer Privacy Act (CCPA), Cal. Civ. Code § 1798.100-1798.199 (2020). <https://oag.ca.gov/privacy/ccpa>

Castells, M. (2000). *The rise of the network society: The information age: Economy, society, and culture* (2nd ed.). Oxford: Blackwell.

Castells, M. (2009). *Communication power*. Oxford: Oxford University Press.

Dubar, C. (2005). *La socialisation: Construction des identités sociales et professionnelles* (4th ed.). Paris: Armand Colin.

European Parliament and Council. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union*, L119, 1-88. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

European Parliament and Council. (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services (Digital Services Act) and amending Directive 2000/31/EC. *Official Journal of the European Union*, L277, 1-68. <https://eur-lex.europa.eu/eli/reg/2022/2065/oj>

Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Stanford: Stanford University Press.

Goffman, E. (1988). *Estigma: Notas sobre a manipulação da identidade deteriorada*. Rio de Janeiro: LTC.

Google. (2024). YouTube Help. Google Support. Recuperado de <https://support.google.com/youtube>

Hall, S. (1996). *Questions of cultural identity*. London: SAGE Publications.

Hamanaka, S. (2020). The role of digital media in the 2011 Egyptian revolution. *Democratization*, 27(5), 777-796. <https://doi.org/10.1080/13510347.2020.1737676>

Huang, S., & Yang, T. (2024). Auditing entertainment traps on YouTube: How do recommendation algorithms pull users away from news. *Political Communication*, 1-19. <https://doi.org/10.1080/10584609.2024.2343769>

ImaRgens. (2024, 9 de agosto). Conheça os 30 filmes curta-metragem resultantes da 2ª oficina especial de audiovisual do NUPEPA-ImaRgens. <https://www.imargens.com.br/single-post/conhe%C3%A7a-os-30-filmes-curta-metragem-resultantes-da-2%C2%AA-oficina-especial-de-audiovisual-do-nupepa-imar>

Naím, M. (2007). Missing links: The YouTube effect. *Foreign Policy*, (158), 103-104. <http://www.jstor.org/stable/25462134>

Nikolov, D., Lalmas, M., Flammini, A., & Menczer, F. (2019). Quantifying biases in online information exposure. *Journal of the Association for Information Science and Technology*, 70(3), 218-229. <https://doi.org/10.1002/asi.24121>

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin Press.

Peralta, A. F., Neri, M., Kertész, J., & Iñiguez, G. (2021). Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions. *Physical Review E*, 104(4), 044312. <https://doi.org/10.1103/PhysRevE.104.044312>

Priante, A., Ehrenhard, M. L., van den Broek, T., & Need, A. (2018). Identity and collective action via computer-mediated communication: A review and agenda for future research. *New Media & Society*, 20(7), 2647-2669. <https://doi.org/10.1177/1461444817731389>

Reuter, C., Stieglitz, S., & Imran, M. (2020). Social media in conflicts and crises. *Behaviour & Information Technology*, 39(3), 241-251. <https://doi.org/10.1080/0144929X.2019.1629025>

Santos, S. F. (2024, agosto 31). Musk's X banned in Brazil after disinformation row. BBC News. Recuperado de <https://www.bbc.com/news/articles/c5y3rnl5qv3o>

Soh, S., Talaifar, S., & Harari, G. M. (2024). Identity development in the digital context. *Social and Personality Psychology Compass*. <https://doi.org/10.1111/spc3.12940>

Spann, B., Agarwal, N., Stafford, D., & Okeke, O. (2023). Evaluating the emergence of collective identity using socio-computational techniques. In *Companion Proceedings of the ACM Web Conference 2023* (pp. 1389-1394). <https://doi.org/10.1145/3543873.3587637>

Tarafdar, M., & Ray, D. K. (2021). Role of social media in social protest cycles: A sociomaterial examination. *Information Systems Research*, 32(3), 1066-1090. <https://doi.org/10.1287/isre.2021.1013>

Wentker, M., & Schneider, C. (2022). And she be like 'Tenemos frijoles en la casa': Code-switching and identity construction on YouTube. *Languages*, 7(3), 219. <https://doi.org/10.3390/languages7030219>

Yun, G. W., Allgayer, S., & Park, S.-Y. (2020). Mind your social media manners: Pseudonymity, imaginary audience, and incivility on Facebook vs. YouTube. *International Journal of Communication*, 14, 21-40. Recuperado de <https://ijoc.org/index.php/ijoc/article/view/12345>

DECLARAÇÃO DE DISPONIBILIDADE DE DADOS DA PESQUISA:

Todo o conjunto de dados de apoio aos resultados deste estudo foi disponibilizado em um repositório do GitHub e pode ser acessado em <https://github.com/nupepa/48ANPOCS>

FINANCIAMENTO:

O presente trabalho foi realizado com apoio da Fundação para a Ciência e Tecnologia (FCT - Portugal) e com apoio do Instituto de Comunicação da NOVA (ICNOVA), Universidade NOVA de Lisboa, Portugal. Códigos de financiamento das bolsas doutorais: 2022.14807.BD (Autor 1) e UI/BD/153755/2022 (Autor 2)

CONTRIBUIÇÃO DAS/DOS AUTORES/AS:

Ferreira, Allan: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration, Supervision, Funding acquisition.

Trevisan, Ana Carolina: Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization.

DECLARAÇÃO DE CONFLITO DE INTERESSE:

Os autores declaram que não há conflito de interesses a mencionar.

MINIBIOGRAFIAS DOS/DAS AUTORAS DO PAPER

Allan Herison Ferreira é doutorando em Ciências da Comunicação na Universidade NOVA de Lisboa, com mestrado em Sociologia pela USP e bacharelado em Ciências Sociais pela USP. Bolsista da FCT afiliado ao ICNOVA, pesquisa "Estratégias identitárias em filmes digitais amadores".

Ana Carolina Trevisan é doutoranda em Ciências da Comunicação na NOVA de Lisboa, mestra em Sociologia pela USP e bacharel em Sociologia e Psicologia. Bolsista da FCT pelo IFILNOVA, pesquisa "Estratégias argumentativas nas redes sociais da extrema direita no Brasil e Portugal".

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença [Creative Commons CC-BY](#).
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.