Vaccine Hesitancy on YouTube:
a Competition between Health and Politics

Yelena Mejova
ISI Foundation
Via Chisola 5, Turin, Italy
yelenamejova@acm.org

Michele Tizzani
Department of Applied Mathematics and Computer Science
Technical University of Denmark
Richard Petersens Plads, 2800, Copenhagen, Denmark
micti@dtu.dk

Abstract—**YouTube has rapidly emerged as a predominant platform for content consumption, effectively displacing conventional media such as television and news outlets. A part of the enormous video stream uploaded to this platform includes health-related content, both from official public health organizations, and from any individual or group that can make an account. The quality of information available on YouTube is a critical point of public health safety, especially when concerning major interventions, such as vaccination. This study differentiates itself from previous efforts of auditing YouTube videos on this topic by conducting a systematic daily collection of posted videos mentioning vaccination for the duration of 3 months. We show that the competition for the public's attention is between public health messaging by institutions and individual educators on one side, and commentators on society and politics on the other, the latest contributing the most to the videos expressing stances against vaccination. Videos opposing vaccination are more likely to mention politicians and publication media such as podcasts, reports, and news analysis, on the other hand, videos in favor are more likely to mention specific diseases or health-related topics. Finally, we find that, at the time of analysis, only 2.7% of the videos have been taken down (by the platform or the channel), despite 20.8% of the collected videos having a vaccination hesitant stance, pointing to a lack of moderation activity for hesitant content. The availability of high-quality information is essential to improve awareness and compliance with public health interventions. Our findings help characterize the public discourse around vaccination on one of the largest media platforms, disentangling the role of the different creators and their stances, and as such, they provide important insights for public health communication policy.**

Index Terms—**vaccination, video, audit, YouTube, science, communication**

I. INTRODUCTION

A January 2025 report by Nielsen has found that streaming platforms accounted for the largest share of all TV viewing formats, and that YouTube is the top used service in the United States [1]. YouTube has been acknowledged as an important resource for science communication as early as 2010, when it was found that the CDC's YouTube channel contributed about 12% of engagement around videos about the ongoing H1N1 influenza epidemic [2]. But even then, 23% of the videos on the topic were found to be misleading. During the COVID-19 pandemic, the platform has announced a blanket ban on anti-vaccine misinformation,[1] but the enforcement remained lax,

and recent audits have found a variety of anti-vaccination content [3], [4]. This may be understandable, since de-platforming channels promoting vaccination hesitancy would cost YouTube millions in profits [5].

Although public health professionals and organizations are increasing their presence on social media platforms, they may be competing with potentially millions of other organizations and individuals for the public's attention.[2] "Influencers" who garner large followings may have audiences larger than the accounts of legitimate public health institutions, and may wield their power to either educate and entertain the public, but also to possibly promote unproven or harmful health advice. Such harmful advice may be targeted specifically to already vulnerable populations, such as those having serious illness or those who have had a negative interaction with physicians or unsuccessful treatments, and who have developed a distrust for mainstream medicine.[3] Since the COVID-19 pandemic, the vaccination debate in particular has been politicized, garnering attention from pundits and commentators having limited expertise in public health or medicine.

In the present study, we shed light on the landscape of the YouTube channels posting content mentioning vaccination by collecting a unique dataset of daily collections of videos over a 3-month period. We use the latest NLP methods to detect the video's stance on the benefits of vaccination, and examine the channels contributing to the videos supporting or opposing vaccination. Examining the metadata and transcripts of 3496 videos, we found that about 21% of these had a stance against vaccination. These videos were much more likely to come from channels associated with individuals than organizations, and from those on the topics of politics and society (as opposed to knowledge or health). Worryingly, several channels self-identifying as "doctors" have posted such videos. Further, they are much more likely to mention politicians and media coverage, unlike videos with a stance favoring vaccines, which are more likely to mention specific illnesses which vaccines alleviate. Finally, we find that only 2.7% of the videos against vaccination have been taken down (by the platform or the poster), indicating a lax moderation regime.

---

[1]https://www.businessinsider.com/non-us-anti-vaxxers-are-still-on-youtube-despite-ban-2021-9

[2]https://www.straitstimes.com/opinion/apple-cider-vinegar-how-social-media-gave-rise-to-fraudulent-wellness-influencers

[3]https://www.apa.org/topics/journalism-facts/false-health-information

II. RELATED WORK

Exposure to low-quality information on internet, and social media in particular, has long been connected with vaccine hesitancy. Respondents to a 2021 survey of 27 EU Member States who favored social media as a source of news were much more likely to be vaccine hesitant [6], while Twitter users in the U.S. who followed politicized news sources were less likely to vaccinate during COVID-19 [7]. Some estimates put the effects of this information into perspective: for instance, it is estimated that, in Canada, COVID-19 misinformation has cost at least 2,800 lives and $300 million in hospital expenses [8]. Thus, estimating the quality of vaccine information online has been an active area of research.

YouTube serves as a significant channel for disseminating information, including health and science topics, offering an accessible alternative to traditional learning [9]. However, the platform's open nature means the validity of health-related content, especially concerning vaccination, can range from reliable to dangerously misleading (see [10] for a literature review). Some of the earliest research on YouTube content around vaccination comes from 2007 [11], which performed an audit of videos mentioning "vaccination" or "immunization". The authors found that 32% of the videos were negative, and 20% were ambiguous about immunization. This trend was confirmed by later studies, highlighting common anti-vaccination arguments including alleged side effects, autism links, pharmaceutical lobbying, and adjuvant concerns [12], [13]. The COVID-19 pandemic intensified this, with significant content contradicting established health organizations [3]. This issue extends beyond YouTube, with anti-vaccine messages spreading across platforms [4], although YouTube's recommendation algorithms may sometimes steer users towards pro-vaccine content [14].

The source of information critically influences its reception. Research distinguishes between professionally generated content (PGC) from entities and content from individual creators [15], [16]. In health communication, signals of authority (e.g., medical credentials, institutional affiliation) are often used [10], though anti-vaccination content may mimic scientific authority through specific terminology [17]. Notably, production quality doesn't always correlate with scientific accuracy or creator credentials [18]. In this study, we focus on the signals of authorities the channels may project, as organizations or as individuals. Note that the claims of accreditation or affiliation may be wholly or partially erroneous. However, it is the projection of authority on which we focus in this study. Also note that the quality of the video may not necessarily correlate with the authoritativeness of its creative. For instance, a detailed analysis of 190 videos of popular scientific YouTube channels found that non-affiliated channels such as *SciShow* and *Veritasium* may have high production values and garner followers in the hundreds of thousands [18].

Because most of the previous literature has used a laborious manual process to assess video content, the labeling of the video's stance on the vaccination or immunization was often done manually [4], [17], [19]–[21]. While machine learning methods like SVMs were explored [14], recent Large Language Models (LLMs) offer a more scalable and efficient approach [22], [23], demonstrating effectiveness in classifying vaccine stances on social media [24]. In this work, we use the latest available technology: the Large Language Models (LLMs), which have been shown to perform well on Natural Language Processing (NLP) tasks [22]. In particular, it has been shown that LLMs provide a faster and more efficient option compared to traditional methods for the task of vaccine stance classification [23], whereas another showed the performance of ChatGPT (`gpt-3.5-turbo-0613` model) to attain accuracy between 0.72 and 0.88, when tested on Twitter and Facebook posts [24]. In this study, we use the latest available ChatGPT model (`gpt-4o-mini`) and manually evaluate its output.

III. DATA & METHODS

A. Data Collection

To achieve the aim of a temporal sample of YouTube videos mentioning vaccination, we devised a daily collection routine, which we have executed since November 20, 2024. The dataset for this study considered all collected data up until February 19, 2025, resulting in 3 months of data. The collection is ongoing for future work. Every day, three actions are taken: search for the previous day's videos, a selection of videos of interest from the search, and crawl additional metadata of the selected videos.

The daily search queries the YouTube Search API[4] for videos posted in the previous 24 hours with one or more of the following keywords: *vaccination, vaccine, antivaxx, antivaxxer, novax, provax*, following the same logic as in [25], [26] (note that the search performs stemming by default and forms of words such as "vaccinations" are also retrieved). The settings are set as follows: region as "US", language as English, order of videos by relevance. Up to 500 video IDs are collected. Using these IDs, their metadata is collected using YouTube Video API call,[5] as well as the transcript using the YouTube Transcript API library.[6] The collected videos are then passed through several filters to select those more likely to be concerning the topic of vaccination, and to have enough metadata for analysis. Specifically, we select videos that have a language detected as English (passing about 32% of the videos), that are between 1 and 60 minutes in length (down to 17% of total), have a transcript (13%), and have at least one keyword in the title and transcript, or if there are no keywords in the title, have at least 3 keywords in the transcript. On average, about 9% of the search results pass these filters and are added to the dataset. We then collect additional metadata about the videos, including that about the channels that posted them. In total, 3496 videos were collected.

---

[4]https://developers.google.com/youtube/v3/docs/search/list
[5]https://developers.google.com/youtube/v3/docs/videos
[6]https://pypi.org/project/youtube-transcript-api/

Category	Score range	# Videos	# Channels	V/C
strongly in favor	[0.5, 1]	1188	880	1.35
in favor	(0, 0.5)	744	585	1.27
neutral	[0]	836	718	1.16
against	(-0.5, 0)	353	285	1.24
strongly against	[-1.0, -0.5]	375	291	1.29

B. Data Annotation

Vaccination Stance: To distinguish between videos that may contain vaccine-hesitant content, we use the most recent LLM technology to label the transcript snippets. Focusing on the content relevant to vaccination, we extract 100 word excerpts from the video transcript around the keywords of interest, and select up to 20 for labeling (excerpts for videos having more than 20 are randomly sampled). The 23 820 excerpts (on average, 6.8 per video) were then annotated using the `gpt-4o-mini` model with the following labels: "in favor of vaccination", "against vaccination", "not in favor or against vaccination". The prompt used for the annotation can be seen in the Appendix A V. Two authors annotated a subset of 50 excerpts, the inter-annotator agreement was at 100% overlap, or Cohen's Kappa of 1. The 84% of human labels matched with those of the LLM, though only two of the mismatches were against vs. in favor, and other mismatches involved the neutral label. Given this performance of the model, we extend its excerpt labels to the videos. To do so, we compute the average score over all excerpts in a video (coding *against* as -1, *in favor* as 1, and *not in favor or against* as 0). We then used a Likert-like scoring for the videos using the following ranges: "strongly in favor" [0.5, 1], "in favor" (0, 0.5), "neutral" [0], "against" (-0.5, 0), and "strongly against" vaccines [-1.0, -0.5]. The distribution of labels and their descriptive statistics can be seen in Table I.

Channel Classification: To further understand the role of the most active content creators, we also classify the YouTube channels with the most videos for each stance. Specifically, for each stance, we select the first 30 channels with the most videos in our dataset and manually annotate them. This threshold of 30 was chosen as it represents the most active creators before entering the long tail of the distribution, where channels contributed far fewer videos each. To define the channel classes, we consider two criteria: the *channel owner* and the *channel category*. In particular, the channel owner can be either an *individual*, an *organization*, or neither (*other*). For the content category we focus on five different classes that are relevant for our study. The definition of the category was extrapolated from the primary focus of the channel, as indicated by the thematic content of the videos produced and the content creators' self-declaration of intent. The categories are the following: *health* for the channels focusing on health-related topics; *news* for all the channel self-declared news

outlets, *opinion* for all the channels not self-declared as official news or information sources, *science* for channels self-declared with a focus on science, *doctor* for channels hosted by self-identified doctors, and *others*.

IV. RESULTS

A. Video Characterization

We begin by considering the topical categories associated with the collected videos, which were labeled with different stances. Figures 1 (a-b) show these relationships for the five most frequent topics in our dataset: Health, Knowledge, Politics, Society, and Television. The plots present complementary views: the distribution of stances within each topic (a), and the distribution of topics within each stance (b). Overall, Society and Health were the most common topics assigned to videos in our collection. Figure 1a displays the distribution of stances within each topic. Videos categorized under Health predominantly exhibit "Strongly In Favor" (SIF) or "In Favor" (IF) stances. On the other hand, topics related to Society and Politics show a markedly higher proportion of "Against" (A) and "Strongly Against" (SA) stances. The Knowledge and Television topics display more intermediate distributions, generally showing lower proportions of hesitant (A and SA) stances compared to Society and Politics. Figure 1b shows the topic distribution within each stance, reinforcing these observations. The SIF and IF stance categories are mainly composed of videos tagged with the Health topic. In contrast, the A and SA stance categories draw heavily from videos tagged with Society and Politics.

Besides the general categories, the videos are also annotated with "tags" which are not standardized. We lower-case these and remove all non-alphanumeric characters. We then compute the odds ratio that a tag appears in a video labeled as SIF and not in SA (low scores of this metric also tell us the reverse). Table II shows the top tags for the two labels. Those associated with the videos labeled as SIF include numerous mentions of diseases and conditions which vaccines prevent, such as *hpv*, *cervical cancer*, *measles*, *human papilloma virus*, *whooping cough*, and *hiv*. It also specifically mentions *women's health* and developing regions such as *africa* and *india*. The US *Centers for Disease Control and Prevention* and *Department of Health & Human Services (HHS)* are also prominent. On the other hand, the tags more likely to appear in the SA videos mention politics (*republicans*, *RFK* in relation to Robert F. Kennedy Jr. who was nominated by the US president as the Secretary of Health and Human Services, and *Elizabeth Warren*, a US senator, as well as *Trump* and *Biden*, both US presidents). A political slogan is also present: *make america healthy again*, echoing the catchphrase made popular by Donald Trump (*Make America Great Again*). Note that no public health institution is mentioned in these, but instead are *reports*, *newsanalysis* and *podcast*.

Next, we stem the tags, group them, and compute the odds ratios (OR) of the tag group appearing in videos labeled as SIF against those labeled SA. Figure 2b shows the OR for the top 10 most frequent grouped tags. We find clear thematic

distinctions: tags related to "science" and "health" are significantly more likely to appear in SIF videos, with odds ratios substantially greater than 1 and confidence intervals entirely above the null value. In contrast, tags associated with "news", "education", "rfk_jk", "covid", "trump", and "politics" all show odds ratios significantly less than 1, indicating these terms were disproportionately more common in SA videos. Tags directly mentioning "vaccine" and "medicine" did not show a statistically significant difference between the groups.

B. Channel Characterization

Next, we focus on the channels posting these videos. We begin by examining the videos with each of the annotated label categories. As can be seen in Table I, the proportion of videos to channels is fairly stable across the categories, at about 1.3 videos per channel. Table III shows the top 10 channels in each category, and for the top 3 – the number of subscribers and an excerpt from their description (full descriptions omitted for brevity). Note that, since the labels are per video, it is possible for a prolific channel to contribute videos to several label categories. For instance, "Forbes Breaking News" is the top contributor of videos labeled as SIF and N. Similarly, a channel of "Dr. John Campbell" is in the list of top contributing channels both in the "against" and "strongly against" categories. Furthermore, the channel "MicrobeTV" contributes videos to both IF and A categories.

In general, we find news organizations in the SIF and IF categories, as well as health-oriented organizations (WHO) and educational efforts (HIV RNA Test Guide). On the other hand, more individuals are visible at the top of the ranking for the SA and A categories (Dr. John Campbell, Russell Brand); however, self-described "news" channels are also present, including "Rebel News" and "GBNews".

We find that the potential reach of these channels varies widely, with individual-centered channels, such as "Dr. John Campbell" having viewership in the millions, whereas, for instance, WHO's channel has smaller (but still substantial) 930K subscribers. Throughout the dataset, we note that there isn't a single channel that dominates the sample: the top contributing channel – "HIV RNA Test Guide" – is responsible for only 0.8% of videos in the dataset, indicating the breadth of the sampling method used in this study.

We also analyzed the distribution of stances across different channel metadata dimensions, as visualized in Figure 3. This includes channel topics (Figure 3 a-b), manually annotated channel content categories (Figures 3 c-d), and channel owner types (Figures 3 e-f). Each pair of plots presents complementary views: the distribution of stances within each metadata category (Figures 3 a, c, e) and the distribution of metadata categories within each stance (Figures 3 b, d, f). We see a thematic heterogeneity based on stance across the channel topics. Videos categorized under the "Health" topic predominantly exhibit SIF or IF stances towards vaccination. Conversely, topics related to "Society" and "Politics" show a higher proportion of A and SA stances. "Knowledge" and "Lifestyle" topics display more intermediate distributions,

generally leaning less towards hesitant stances. The inverse perspective, shown in Figure 3b, confirms this pattern: SIF and IF stances are disproportionately composed of "Health" related content, while A and SA stances draw heavily from videos tagged with "Society" and "Politics". Finally, we examined the odds ratios for the top 5 most frequent channel topics, comparing videos classified as SIF against those classified as SA, with SA serving as the reference group (Figure 2c). A clear thematic distinction emerged. Channels categorized as 'Health' showed the strongest positive association with the SIF label (OR appearing visually around 2.5), followed by "Lifestyle" (OR $\simeq$ 2.0) and "Knowledge" (OR $\simeq$ 1.8). On the other hand, channels focusing on "Society" (OR $\simeq$ 0.7) and "Politics" (OR $\simeq$ 0.5) were significantly less likely to be associated with SIF content. These findings highlight that channels covering societal and political discourse are disproportionately linked to content strongly opposing vaccination, whereas channels focused on health, lifestyle, and general knowledge predominantly align with pro-vaccination stances.

Analysis of manually annotated channel categories (Figure 3c) highlights that the "Opinion" category is greatly associated with A and SA stances. In contrast, channels categorized as "Health" and "Science" predominantly feature SIF and IF content. The "News" category displays a more mixed distribution, including substantial proportions of neutral (N), IF, and SIF stances, but also a noticeable share of A stances. Interestingly, the "Doctor" category, while featuring significant favorable content, also contains a high proportion of N, A, and SA stances, suggesting creators identified as doctors are not uniformly pro-vaccination in this sample. Figure 3d reinforces these findings, showing the "Opinion" category's dominance within A and SA stances, while "Health", "News", and "Science" are the primary contributors to SIF and IF stances.

Examining the type of channel owner (Figure 3e), videos posted by "Individual" creators exhibit a considerably larger proportion of A and SA stances compared to those from "Organization" or "Other" types. "Organization"-owned channels show a strong tendency towards SIF and IF stances. Figure 3f demonstrates that while "Organization" owners contribute significantly to SIF and IF content, "Individual" creators dominate the A and SA stance categories and also represent a substantial portion of the favorable stance video

These distributions highlight clear associations between video stance and the topic, declared content category, and owner type of the channel, suggesting systematic differences in how vaccination is framed across different types of YouTube content creators and subject areas.

C. Platform Moderation

When the metadata of the videos was re-crawled at the end of the collection, we found that 96 videos failed to return results. We manually examined the errors the YouTube interface provided for these, and found that for 43, the platform simply tells the users that the "This video isn't available anymore", 24 have been made private by the uploader, 15 were
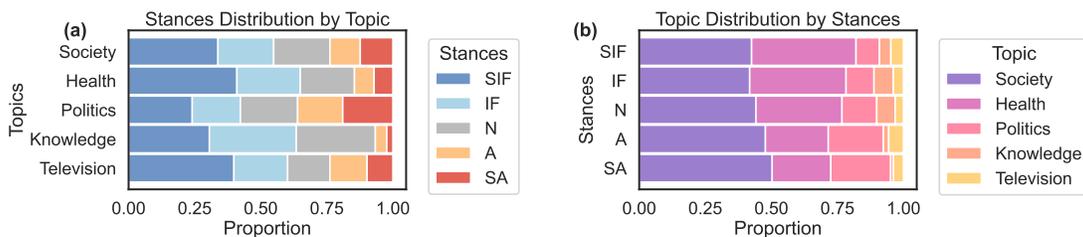
Fig. 1. The stacked bar plots show the stance distributions by topic (a) and topic distribution by stances (b). The left plot has topics on the y-axis and proportion on the x-axis. The right plot has stances on the y-axis and proportions on the x-axis.

TABLE II
TOP TAGS OF VIDEOS BY ODDS RATIO (SHOWN IN PARENTHESES): ON THE LEFT, THE TAGS THAT ARE MORE LIKELY TO APPEAR IN VIDEOS LABELED AS "STRONGLY IN FAVOR", ON THE RIGHT, THOSE MORE LIKELY TO APPEAR IN VIDEOS THAT ARE "STRONGLY AGAINST" VACCINES.

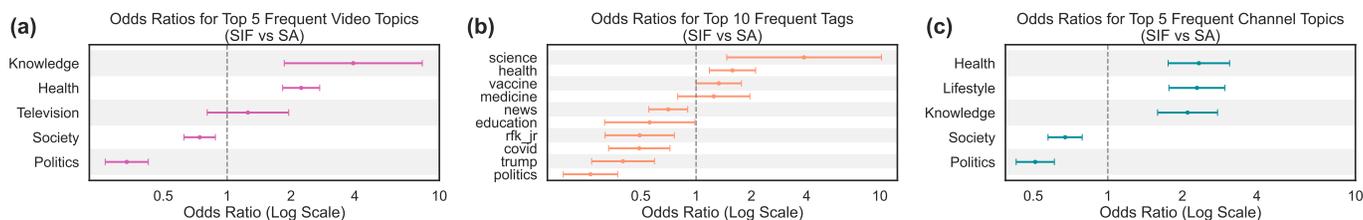| Strongly in Favor	Strongly Against
local (0.02), cervicalcancer (0.04), hpv (0.04), womenshealth (0.05), measles (0.06), centersfordiseasecontrolandprevention (0.06), rsv (0.06), humanpapillomavirus (0.07), vaccinedevelopment (0.07), whoopingcough (0.07), currentaffairs (0.08), africa (0.08), howvaccineswork (0.08), cbs (0.09), hivvaccine (0.09), vaccinescience (0.09), hiv (0.09), smallpox (0.09), india (0.09), hhs (0.09)	reports (5.05), disease (4.34), uk (4.23), newsanalysis (3.45), covid-vaccine (3.10), republicans (2.92), biology (2.72), podcast (2.64), rfk (2.60), antivaccine (2.39), elizabethwarren (2.39), makeamericahealthyagain (2.39), businessnews (2.39), republican (2.27), trump (1.86), medicine (1.86), biden (1.86), election (1.86), economy (1.86), senateconfirmationhearing (1.86)

Fig. 2. Odds Ratios (ORs) with 95% Confidence Intervals (CIs), comparing features associated with videos classified as Strongly In Favor (SIF) versus Strongly Against (SA) vaccination, with SA as the reference group. An OR > 1: feature is more associated with the SIF stance, OR < 1: with the SA stance. The panels show results for: (a) the top 5 frequent video topics, (b) the top 10 frequent grouped video tags, and (c) the top 5 frequent channel topics.

unavailable due to the posting account having been terminated by the platform, 8 were removed by the uploader (a message distinct from the video simply being unavailable), 4 were removed due to violating YouTube's Community Guidelines, one because the uploader has closed their YouTube account, and one because of a copyright claim. We consider the video disappearances attributed to the author separately from all others (which we attribute to the platform), and find rates shown in Figure 4. We find that, although the removals by the author of the video have about the same rate across the categories, videos labeled as A or SA were much more likely to be removed by the platform.

Considering the error codes more closely, we find that all videos removed due to violating YouTube's Community Guidelines were either labeled as A or SA. On the other hand, videos that belonged to channels terminated by the platform span all five categories. Whenever the video was made private by its author, the label is more likely to be IF or SIF (65% of such videos) than A or SA (33%). Thus, we find a variety of reasons for the disappearance of videos, with some evidence of the platform's policies affecting more the vaccine-skeptic

side when it comes to sanctioning content, but not necessarily the accounts themselves.

V. DISCUSSION

In this work, we analyze longitudinal data of vaccine-related content on YouTube, a platform that has become a dominant source of information and entertainment for a vast portion of the digital population. Our analysis shows a clear distinction in the thematic focus of the videos based on their stance on vaccination and is visually summarized in Figures 1 and 3. While pro-vaccine content predominantly centers around "Health" and specific diseases, those expressing hesitancy are more likely to fall under "Society" and "Politics". The analysis of video tags provides further evidence of the thematic divergence. Pro-vaccine videos focus on specific diseases and health organizations, while anti-vaccine videos frequently mention political figures, slogans, and terms like "news analysis" and "reports". This finding underlines the difficulty in discerning health information from broader socio-political commentary, making it difficult for the user to identify credible sources. This

Channel	Subscribers	Description excerpt
Strongly in favor of vaccines (SIF)		
Forbes Breaking News (12)	4.64M	[none]
HIV RNA Test Guide (8)	5.14K	a platform where people can learn about common sexually transmitted diseases i.e STDs.
World Health Organization (WHO) (7)	930K	The official public health information Youtube channel of the World Health Organization (WHO)
Others: Dr Kat, Epidemiologist (7), Citizen TV Kenya (7), TTT Live Online (6), CP24 (6), 9NEWS (6), GoodRx (5), Channels Television (5)		
In favor of vaccines (IF)		
HIV RNA Test Guide (19)	5.14K	a platform where people can learn about common sexually transmitted diseases i.e, STDs.
Dawah Institute (9)	3.53K	dawahinstitute.org
MicrobeTV (9)	135K	I'm Vincent Racaniello, Earth's virology Professor and I believe that education should be free.
Others: CreakyJoints (7), Citizen TV Kenya (6), Forbes Breaking News (5), Casina Pio IV (5), WION (5), Medscape (5), KGW News (4)		
Neutral of vaccines (N)		
Forbes Breaking News (7)	4.64M	[none]
Vejon Health (5)	157K	At Vejon Health, we want to make the world a better place for you and future generations by improving the understanding and management of complex diseases.
Stuff (4)	68K	Stuff is New Zealand's largest and most popular news site.
Others: Dr Neha Taneja's Community Medicine (4), WION (4), NTV Kenya (4), HIV AIDS PREVENTION (4), Chuck Miller Plays (4), KBC Channel 1 (3), Kenya Digital News (3)		
Against of vaccines (A)		
Vejon Health (9)	157K	At Vejon Health, we want to make the world a better place for you and future generations by improving the understanding and management of complex diseases.
Dr. John Campbell (8)	3.19M	My name is John Campbell and I am a retired Nurse Teacher and former clinical nurse based in England. [...] Disclaimer; These media including videos, book, e-books, articles, podcasts are not peer-reviewed.
Russell Brand (8)	6.84M	Something ancient and original is Awakening within us. Here we can learn from elders and grow together in Good Faith, flawed but rising together.
Others: The Jimmy Dore Show (5), GBNews (4), Sara Gonzales Unfiltered (4), TalkTV (4), Rebel News (4), The Next News Network (3), MicrobeTV (3)		
Strongly against of vaccines (SA)		
Find the Best (12)	2.77K	At Find the Best, we are dedicated to helping you make informed decisions by providing comprehensive information on equipment, companies, products, and services.
GBNews (9)	1.75M	Keep up to date with GB News at gbnews.com or on X GBNEWS
Dr. John Campbell (9)	3.19M	My name is John Campbell and I am a retired Nurse Teacher and former clinical nurse based in England. [...] Disclaimer; These media including videos, books, e-books, articles, and podcasts, are not peer-reviewed.
Others: The Jimmy Dore Show (8), Redacted (7), Russell Brand (6), Kim Iversen (5), Rebel News (5), Viva Frei (3), Starkey Farmstead (3)		

aligns with previous results suggesting the strong politicization of health issues, particularly during COVID-19 [4].

When analyzing the role of individual creators and organizations (Figures 3e-f), we discover that, perhaps unsurprisingly, established news organizations and health institutions (e.g., WHO and Forbes) are major contributors to pro-vaccine content, while individual creators with substantial followings are prominent in the "against" and "strongly against" content. This aligns with previous studies showing individual users "content can be more popular even if less scientifically accurate" [27].

An analogous picture emerges when examining the channel category (Figure 3e-f). Except for the health and science channels, which predominantly support vaccination, more heterogeneous stances characterize the remaining categories. The "News" category shows a more mixed distribution, with a substantial proportion of videos in the "neutral", "in favor," and "strongly in favor" categories, but also a noticeable presence of "against" videos. Online news outlets might be less subject to the reliability of sources than traditional ones, reinforcing the need for users to critically evaluate the credibility of news channels. Surprisingly, the "doctor" category doesn't fully align with the "health" category, falling more into the hesitant stances than the favoring ones. If the self-description is accurate, this could point to the personal positions of healthcare workers. However, the creators' self-declared certifications are difficult to verify. Besides, the strong association between the "opinion" category and anti-vaccine stances indicates that a significant portion of anti-vaccine content on YouTube is not presented as factual news or scientific information, but rather as personal opinions, commentary, or analysis. This critical distinction suggests that much of the anti-vaccine content is not attempting to *appear* as objective reporting, but it's explicitly subjective. The prevalence of the "Society" topic in hesitant content suggests that opposition to vaccination is often framed not as a purely scientific or medical issue, but as a broader societal concern, potentially touching on issues of individual liberty, government overreach, distrust of institutions, or cultural values. This framing allows anti-vaccine narratives to tap into pre-existing social and political anxieties, making them more persuasive and resistant to purely
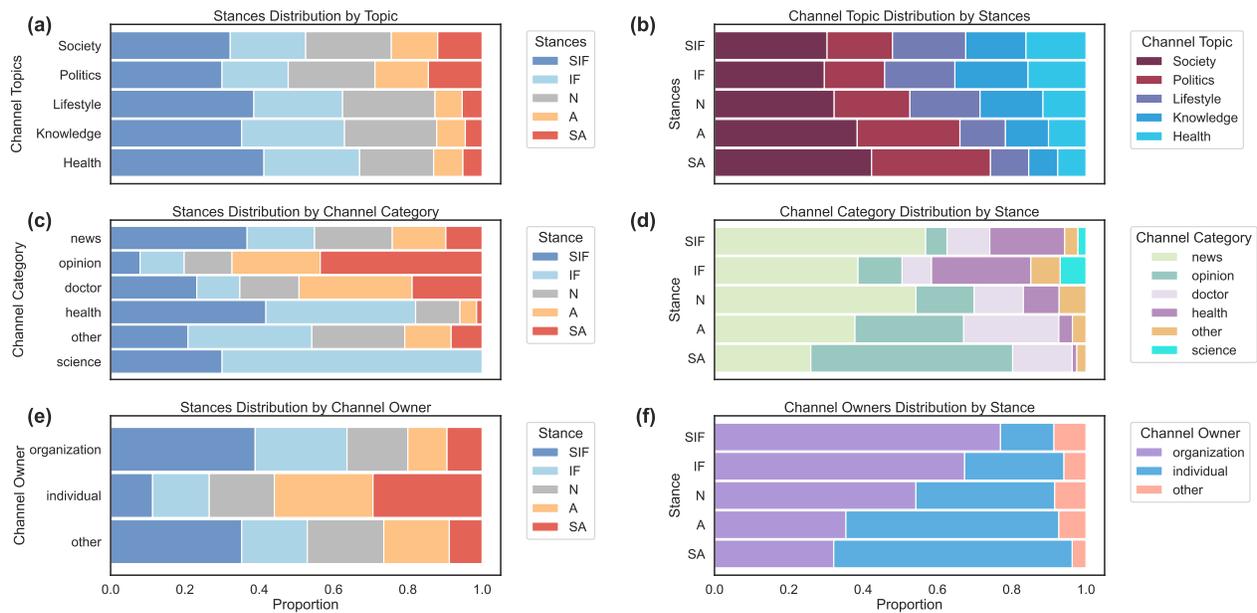
Fig. 3. Channel characterization. Distribution of stances across different topics (a), channel categories (c), and different channel owners (e). Distribution of topics (b), channel categories (d), and channel owner types (f) across stances. The charts are row-normalized (a, c, e) and column-normalized (b, d, f). (a) shows the proportion of each stance (SIF: Strongly In Favor, IF: In Favor, N: Neutral, A: Against, SA: Strongly Against) within each topic. (b) shows the proportion of each channel owner category within each stance. (c) shows the proportion of each stance within each content category. (d) shows the proportion of each content category within each stance. (e) shows the proportion of each stance within the channel owner category. (f) shows the proportion of each channel owner category within each stance.
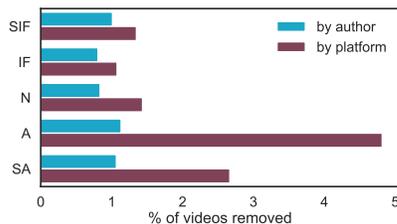

Fig. 4. Percentage of videos in each category that have been removed either by the poster of the video, or by the platform.

factual rebuttals. Videos categorized as "Knowledge" are less likely to be hesitant. This suggests that videos focused on general knowledge dissemination, even if not directly related to health, tend to avoid promoting vaccine hesitancy. The "Lifestyle" category is more evenly distributed, but still leans towards pro-vaccine stances. This could indicate that content creators focused on lifestyle topics (e.g., personal vlogs, daily routines) are less likely to engage in politically charged debates about vaccination, or may even subtly promote positive health behaviors.

Lastly, despite YouTube's strong policies against anti-vaccination content (especially post-COVID),[7] our findings indicate a relatively low rate of video takedowns (2.7%) even for content classified as "strongly against" vaccination (20.8% of the dataset). While we found evidence of platform intervention (videos removed for violating Community Guidelines

[7] https://www.bbc.com/news/technology-58743252

were predominantly anti-vaccine), further labeling the data for misinformation will reveal the coverage of the platform's moderation efforts.

Limitations. This study, in its three-month duration, offers only a snapshot of a rapidly evolving discourse around vaccination. Keyword-based collection might miss coded or emerging language, while LLM stance classification, despite validation, can struggle with distinguishing more complex stances. Although the scope of the study was to capture content that requires a longer user attention span, and we selected only English videos to better design the classification task, the exclusion of the YouTube Shorts format and the focus primarily one language further restricts the scope and generalizability.

Future work. Future work should prioritize longitudinal studies over longer periods to capture evolving narratives and event impacts. Investigating the interplay between video context (style, tone), stance, and user engagement metrics (views, likes, and comments) is crucial, particularly to compare the relative influence of individual creators versus organizations. Dedicated analysis of Shorts and expansion to diverse languages, regions, and platforms will also make the analysis extendable to multiple contexts.

Ethical concerns. Key ethical considerations include the potential risks associated with identifying public channels, even without full anonymization, which could expose creators to unwanted attention. The research involves sensitive interpretation and framing of results, and responsible conduct requires acknowledging the broader implications of large-scale analysis

of public online expression and ensuring findings are used constructively to support public health communication.

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

REFERENCES

[1] Nielsen, "Explore how americans watch tv across platforms," 2025. [Online]. Available: https://www.nielsen.com/data-center/the-gauge/ 1

[2] A. Pandey, N. Patni, M. Singh, A. Sood, and G. Singh, "Youtube as a source of information on the h1n1 influenza pandemic," *American journal of preventive medicine*, vol. 38, no. 3, pp. e1–e3, 2010. 1

[3] H. O.-Y. Li, E. Pastukhova, O. Brandts-Longtin, M. G. Tan, and M. G. Kirchhof, "Youtube as a source of misinformation on covid-19 vaccination: a systematic analysis," *BMJ global health*, vol. 7, no. 3, p. e008334, 2022. 1, 2

[4] A. Gruzd, D. Abul-Fottouh, M. Y. Song, and A. Saiphoo, "From Facebook to YouTube: The potential exposure to COVID-19 anti-vaccine videos on social media," *Social media+ society*, vol. 9, no. 1, p. 20563051221150403, 2023, publisher: SAGE Publications Sage UK: London, England. 1, 2, 6

[5] Center for Countering Digital Hate, "Pandemic profiteers. the business of anti-vaxx," 6 2021. [Online]. Available: https://counterhate.com/research/pandemic-profiteers/ 1

[6] M. Mascherini and S. Nivakoski, "Social media use and vaccine hesitancy in the european union," *Vaccine*, vol. 40, no. 14, pp. 2215–2225, 2022. 2

[7] S. Rathje, J. K. He, J. Roozenbeek, J. J. Van Bavel, and S. Van Der Linden, "Social media behavior is associated with vaccine hesitancy," *PNAS nexus*, vol. 1, no. 4, p. pgac207, 2022. 2

[8] D. Major, "Covid-19 misinformation cost at least 2,800 lives and $300m, new report says," 2023. [Online]. Available: https://www.cbc.ca/news/politics/cost-of-covid-19-misinformation-study-1.6726356 2

[9] S. Kohler and T. C. Dietrich, "Potentials and limitations of educational videos on YouTube for science communication," *Frontiers in communication*, vol. 6, p. 581302, 2021, publisher: Frontiers Media SA. 2

[10] K. Haslam, H. Doucette, S. Hachey, T. MacCallum, D. Zwicker, M. Smith-Brilliant, and R. Gilbert, "YouTube videos as health decision aids for the public: an integrative review," *Canadian Journal of Dental Hygiene*, vol. 53, no. 1, p. 53, 2019, publisher: Canadian Dental Hygienists Association. 2

[11] J. Keelan, V. Pavri-Garcia, G. Tomlinson, and K. Wilson, "Youtube as a source of information on immunization: a content analysis," *jama*, vol. 298, no. 21, pp. 2482–2484, 2007. 2

[12] G. Donzelli, G. Palomba, I. Federigi, F. Aquino, L. Cioni, M. Verani, A. Carducci, and P. Lopalco, "Misinformation on vaccination: A quantitative analysis of youtube videos," *Human vaccines & immunotherapeutics*, vol. 14, no. 7, pp. 1654–1659, 2018. 2

[13] M. Lahouati, A. De Coucy, J. Sarlangue, and C. Cazanave, "Spread of vaccine hesitancy in france: What about youtube™?" *Vaccine*, vol. 38, no. 36, pp. 5779–5782, 2020. 2

[14] Y. M. M. Ng, K. Hoffmann Pham, and M. Luengo-Oroz, "Exploring YouTube's recommendation system in the context of COVID-19 vaccines: Computational and comparative analysis of video trajectories," *Journal of medical Internet research*, vol. 25, p. e49061, 2023, publisher: JMIR Publications Toronto, Canada. 2

[15] D. J. Welbourne and W. J. Grant, "Science communication on YouTube: Factors that affect channel and video popularity," *Public understanding of science*, vol. 25, no. 6, pp. 706–718, 2016, publisher: SAGE Publications Sage UK: London, England. 2

[16] J. Kim, "The institutionalization of youtube: From user-generated content to professionally generated content," *Media, culture & society*, vol. 34, no. 1, pp. 53–67, 2012. 2

[17] N. Yiannakoulias, C. E. Slavik, and M. Chase, "Expressions of pro- and anti-vaccine sentiment on youtube," *Vaccine*, vol. 37, no. 15, pp. 2057–2064, 2019. 2

[18] J. Munoz Morcillo, K. Czurda, and C. Y. Robertson-von Trotha, "Typologies of the popular science web video," *J. Sci. Commun.*, vol. 15, no. 04, p. A02, May 2016. 2

[19] J. Allgaier, "Science and environmental communication on YouTube: Strategically distorted communications in online videos on climate change and climate engineering," *Frontiers in communication*, vol. 4, p. 36, 2019, publisher: Frontiers Media SA. 2

[20] M. Ledwich and A. Zaitsev, "Algorithmic extremism: Examining YouTube's rabbit hole of radicalization," *arXiv preprint arXiv:1912.11211*, 2019. 2

[21] R. Briones, X. Nan, K. Madden, and L. Waks, "When vaccines go viral: an analysis of hpv vaccine coverage on youtube," *Health communication*, vol. 27, no. 5, pp. 478–485, 2012. 2

[22] J. Fields, K. Chovanec, and P. Madiraju, "A survey of text classification with transformers: How wide? how large? how long? how accurate? how expensive? how safe?" *IEEE Access*, vol. 12, pp. 6518–6531, 2024. 2

[23] L. Espinosa and M. Salathé, "Use of large language models as a scalable approach to understanding public health discourse," *PLOS Digital Health*, vol. 3, no. 10, pp. 1–15, 10 2024. [Online]. Available: https://doi.org/10.1371/journal.pdig.0000631 2

[24] S. Kim, K. Kim, and C. Wonjeong Jo, "Accuracy of a large language model in distinguishing anti- and pro-vaccination messages on social media: The case of human papillomavirus vaccination," *Preventive Medicine Reports*, vol. 42, p. 102723, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2211335524001384 2

[25] G. Crupi, Y. Mejova, M. Tizzani, D. Paolotti, and A. Panisson, "Echoes through Time: Evolution of the Italian COVID-19 Vaccination Debate," *Proceedings of the International AAAI Conference on Web and Social Media*, vol. 16, pp. 102–113, May 2022. [Online]. Available: https://ojs.aaai.org/index.php/ICWSM/article/view/19276 2

[26] G. Paoletti, L. Dall'Amico, K. Kalimeri, J. Lenti, Y. Mejova, D. Paolotti, M. Starnini, and M. Tizzani, "Political context of the european vaccine debate on twitter," *Sci. Rep.*, vol. 14, no. 1, p. 4397, Feb. 2024. 2

[27] D. J. Welbourne and W. J. Grant, "Science communication on YouTube: Factors that affect channel and video popularity," *Public Understanding of Science*, vol. 25, no. 6, pp. 706–718, Aug. 2016, publisher: SAGE Publications Ltd. [Online]. Available: https://doi.org/10.1177/0963662515572068 6

APPENDIX

Prompt for Stance Labeling

The prompt to ChatGPT model `gpt-4o-mini`:

Your task is to determine whether the provided texts are in favor of vaccination, against vaccination, or neither.

For the text, assign one of the following labels based on your assessment:

1. "in favor of vaccination" if the overall message is that vaccines are good and helpful, and people should vaccinate. Alternatively, the text could quote or talk about somebody who is against vaccines, but if the text is against that person, the label should be "in favor of vaccination".

2. "against vaccination" if the overall message is that vaccines are dangerous or harmful and people should not vaccinate or given a choice not to vaccinate. Alternatively, the text could quote or talk about somebody who is in favor of vaccines, but if the text is against that person, the label should be "against vaccination".

3. "neither" if the text is not in favor or against vaccines.

Reply providing just the label.

Here is the text: *[text]*