



# Coronavirus News and Information on YouTube: A Content Analysis of Popular Search Terms

Nahema Marchal, Hubert Au, Philip N. Howard

## SUMMARY

Social media is reshaping how people access information about the coronavirus crisis. The online video-sharing website YouTube has emerged as a major purveyor of health and wellbeing information. In this memo, we examine the nature and structural properties of a sample of 320 YouTube videos related to the coronavirus outbreak. We find that:

- four-fifths of the channels sharing coronavirus news and information are maintained by professional news outlets and that the channels of public health agencies are rarely, if ever, returned with search results;
- searches for popular coronavirus-related terms return mostly factual and neutral video results, with low volumes of conspiratorial or junk science video results;
- highly politicized health news and information receives on average more public engagement in the form of comments than any other type of videos.

## INTRODUCTION

The recent outbreak of the coronavirus disease, known as COVID-19, is an ongoing public health emergency that first emerged in Wuhan, China in December 2019. As of April 17th, 2020, nearly 2,113,230 cases of infections have been recorded in a total of 210 countries worldwide.<sup>[1]</sup> Since the outbreak's onset, misinformation about the virus, its provenance and scale have proliferated online.<sup>[2]</sup> On February 2<sup>nd</sup>, the World Health Organization (WHO) director warned against the public health consequences of an “infodemic”—an abundance of potentially inaccurate claims—that would make it difficult for citizens to find reliable guidance when needed.<sup>[3]</sup>

With over 2 billion monthly active users, the online video-sharing website YouTube has emerged as a major source of information about science, technology and health in recent years, especially for young people.<sup>[4]</sup> In the UK alone, an estimated 35.6 million people visit the site each month, making it the most popular online video platform in the country. The average UK adult spends around half an hour on YouTube every day and screen time has been increasing during quarantine.<sup>[5],[6]</sup>

In the context of a major public health crisis such as the 2019 coronavirus outbreak, it is therefore important to understand the type and quality of content that users are likely to encounter when searching for information about the virus on YouTube. In this memo, we ask: Which sources and channels are most represented in search results? To what extent is video content around

the coronavirus's origin, transmission, and cure being politicized, and how much of it is factually inaccurate, misleading, or conspiratorial? How much public engagement is received by different types of videos?

## DATA AND METHOD

The goal of this study is to provide a snapshot of the type of content average users are likely to come across when searching for information about the coronavirus on YouTube. YouTube is one of the largest search engines in the world by search volume, second only to Google, and a gateway to the news for a large number of its users. Like many other social media platforms, YouTube is powered by a search algorithm and recommender system based on the principle of “collaborative filtering,” designed to help users navigate the millions of pieces of content available on its site.<sup>[7]</sup> Video creators also actively attempt to optimize their content for search engines.<sup>[8]</sup> Research suggests that YouTube users rarely scroll past the top twenty results when searching for content and that recommendations are responsible for 70% of time spent on the site.<sup>[9],[10]</sup> In an effort to replicate this user experience, for our analysis we thus focus on the first twenty video results returned for a search query, as well as what YouTube calls the most frequent “related” videos displayed in the sidebar (see Online Supplement for full methods specification).

For this study, we perform a content analysis of the top English-language videos associated with four popular coronavirus-related search terms in the UK: “coronavirus UK”, “coronavirus China”, “coronavirus

symptoms” and “coronavirus conspiracy”. We analyze the top twenty videos results provided by YouTube in order of “relevance”, as well as the top sixty related videos for each search term. Our final sample thus consists of 320 videos, eighty per search term.

Using Google Trends, our team first collected a list of the ten most popular search queries entered into the YouTube search bar in the United Kingdom related to “coronavirus” since January 2020. For comparison, we also compiled a list of auto-complete suggestions for the term “coronavirus” in the YouTube GB search bar, using a Google Chrome browser in Incognito Mode (see Online Supplement for complete list). The term “coronavirus” was selected over “COVID-19” as it was by far the most popular search term at the time of data collection. Four search terms that overlapped between both approaches—namely “coronavirus UK”, “coronavirus China”, “coronavirus symptoms”, and “coronavirus conspiracy”—were selected for the final analysis, as these cover a wide range of topics reflecting strong public interest.

Next, we queried the YouTube API’s search function with our four search terms using [YouTube Data Tools \(YTDT\)](#) on March 20<sup>th</sup>, 2020. This generated a list of the top fifty video results for each search term, ranked by YouTube in order of “relevance”—the default setting for this parameter.<sup>[11]</sup> This seed list was then used to further crawl the YouTube network and collect all related videos associated with these top fifty results. Using this technique, we gathered a total of 8,912 videos. Following previous research in this area, our team created four directed network graphs, where each node in the graph corresponded to a video and each edge denoted a connection from a seed video to a related video.<sup>[12]</sup> We then calculated corresponding network statistics to determine which related videos in each network were most frequently suggested in the side bar. Top related videos were filtered by highest in-degree—the number of times a video was featured on the related list of another video. Our final sample consists of 320 videos: the twenty top video results and sixty top related videos associated with each of four search terms.

First, we classified all channels in this sample, according to the following grounded typology:

- **Government & Public Agencies.** Official channels of government agencies and international bodies, such as the US White House or United Nations.
- **Independent Content Creator.** Channels of independent content creators, such as independent vloggers, media commentators, health professionals and educators.
- **Professional Health.** Channels of domestic and intergovernmental public health agencies, hospitals and professional health websites, such as the WHO, the NHS, or WebMD.
- **Professional News.** Channels of established news organizations, broadcasters, digital and print media outlets.
- **State-Backed Media.** Channels of media organizations that are either directly funded by the state and are editorially controlled by their respective governments.

Each video in our sample was then reviewed in-depth and classified into one of the four categories below. A heuristic approach was taken to content classification based on the type of evidence marshalled, degree of politicization and factual accuracy of the information presented:

- **Factual and Neutral.** Factual and high-quality reporting on the coronavirus pandemic. This includes factoids, real-time case trackers and news reports from professional news organizations and public health agencies.
- **Junk and Conspiratorial.** Videos relaying verifiably false information or conspiracy theories about the origin, transmission and treatment of the coronavirus; trutherism; xenophobia and denial of mainstream scientific positions as assessed against WHO public advisory information.
- **Personal and Investigative.** Discussion of the coronavirus pandemic from a personal, testimonial or investigative point of view. This includes videos of patients describing their symptoms, testimonies, recommendations and debunking efforts from independent vloggers and health practitioners, as well as undercover investigations and talk shows.
- **Political.** Videos in which the coronavirus pandemic, its provenance, spread as well as the efficacy of various government responses, is discussed or debated from a political perspective. This includes political comedy shows, debates, podcasts of a political nature, and ideologically motivated fact-checks.
- **Not Applicable (N/A).** Videos unrelated to the coronavirus pandemic.

All videos and associated channels in our sample were manually classified by a team of two coders who achieved high inter-rater reliability following two rounds of rigorous training and test coding (Krippendorff’s  $\alpha = 0.803$  for content codes). Any disagreements were resolved by the authors. A further thematic analysis was conducted to determine the main topics covered in each video (see Online Supplement).

## FINDINGS

### Content and channel types

We found that the vast majority of channels were maintained by professional news organizations (80%) followed by independent content creators (16%), which

include independent healthcare providers, educators, and analysts. By contrast, professional health sites and public health agencies such as the NHS and the WHO constituted less than a single percentage point in this sample (0.3%). A small but non-negligible proportion of the videos analyzed by our team originated from state-backed media outlets (2.2%) including China’s CGTN and GBTimes, as well as Russia Today’s Ruptly. These comprised 6.25% of top and most recommended videos for the query “coronavirus conspiracy” (see Table 1 in Online Supplement).

Table 1 shows that factual and balanced reporting dominates video results when users search for “coronavirus UK” — fully 80% of top twenty results returned are factual or neutral. The picture is more mixed when searching for videos related to coronavirus conspiracy theories, with about one third of the first twenty videos displayed on the results page consisting of politicized content or ideologically-motivated debunking efforts. Interestingly, half of the top video results shown for “coronavirus symptoms” consist of explainers and professional advice delivered by independent vloggers and healthcare professional, testimonial evidence, such as talk shows featuring patients describing their symptoms. Videos peddling junk, factually inaccurate or conspiratorial information are low across all searches, except for “coronavirus China,” where they make up 15% of the top video results displayed.

Trends are broadly similar for related video content (see Table 2). Among the sixty top related videos per search term, only three re-directed to questionable content that runs against public advisory information from the WHO. Investigative reporting and informational videos showcasing personal narratives and testimonial evidence account for a large proportion of the top related videos in the network—22% for “coronavirus conspiracy” and 17% for “coronavirus symptoms” respectively. A user searching for coronavirus news and information related to China, however, is more likely to come across politicized content than through any other search query tested here.

Having classified all videos in our sample, we examined the distribution of video content per channel type. Figure 1 suggests that professional news channels drive most of the factual and journalistic reporting on the coronavirus pandemic and that these predominantly fall under the “News & Politics” label. Public health agencies and professional healthcare are not represented in this figure as the only channel our team coded as such linked to a non-English language video. The Online Supplement presents this data in tabular form. Politicized content is evenly split between professional news channels, on the one hand, and independent content creators as well as state-backed media on the other hand, covering a wide variety of channel types, from “Comedy” to “Music”, “Style” and even “Travel” (See Figure 2 in Online Supplement).

**Table 1. Distribution of top 20 video results, per search query (Percent)**

Type of Content	Search Terms % (N)			
	UK	China	Symptoms	Conspiracy
Factual & neutral	80 (16)	45 (9)	25 (5)	20 (4)
Junk & conspiratorial		15 (3)		5 (1)
Personal & investigative	15 (3)		50 (10)	40 (8)
Political	5 (1)	35 (7)	10 (2)	30 (6)
Non-English		5 (1)	15 (3)	5 (1)
Total	100 (20)	100 (20)	100 (20)	100 (20)

Source: Authors’ calculations based on data collected on 20/03/2020.

Note: Categories are mutually exclusive and columns sum to 100%.

**Table 2. Distribution of top 60 most recommended videos, per search query (Percent)**

Type of Content	Search Terms % (N)			
	UK	China	Symptoms	Conspiracy
Factual & neutral	57 (34)	45 (27)	47 (28)	40 (24)
Junk & conspiratorial				5 (3)
Personal & investigative	17 (10)	13 (8)	17 (10)	22 (13)
Political	18 (11)	28 (17)	15 (9)	15 (9)
Non-English	2 (1)	8 (5)	13 (8)	7 (4)
N/A	7 (4)	5 (3)	8 (5)	12 (7)
Total	100 (60)	100 (60)	100 (60)	100 (60)

Source: Authors’ calculations based on data collected on 20/03/2020.

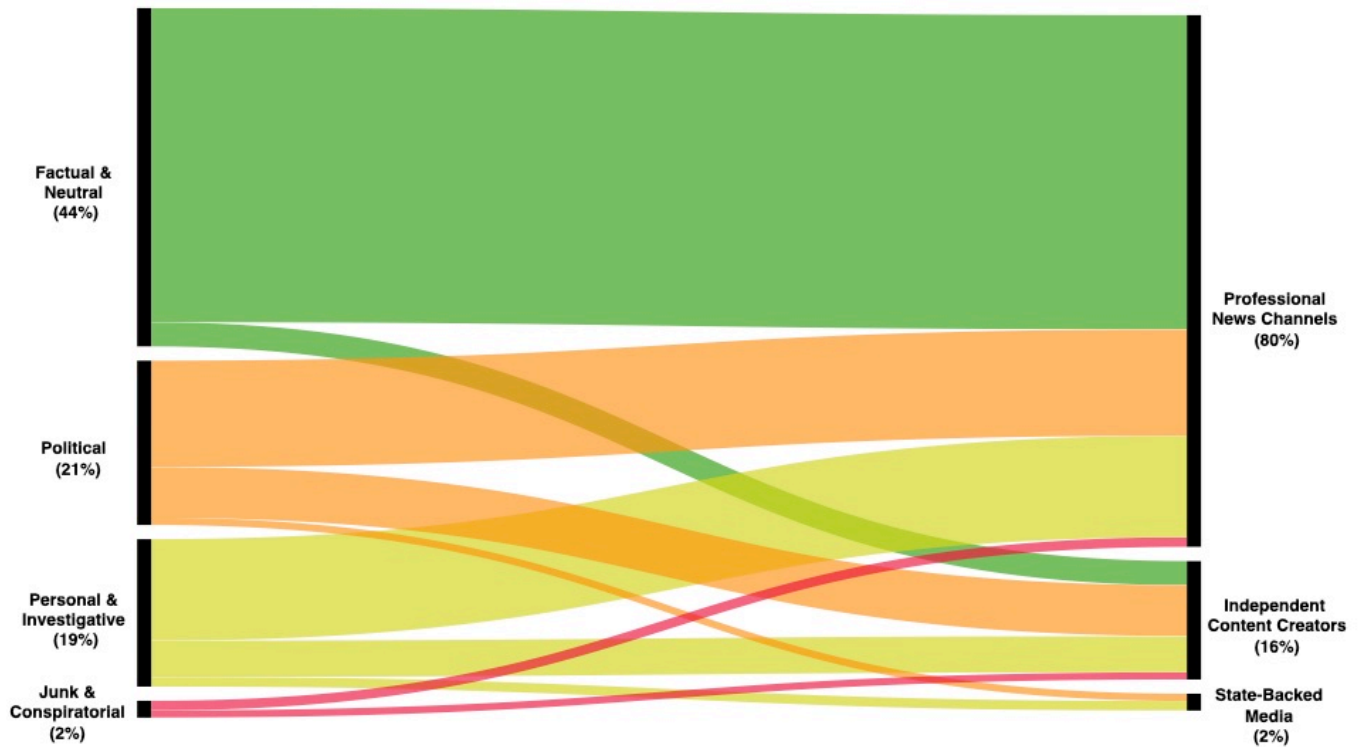
Note: Categories are mutually exclusive and numbers are rounded up to the nearest whole number.

**Table 3. Average engagement with each type of video**

Type of Content	Avg. number of views (million)	Avg. number of comments*	Comment-to-views ratio (1,000 :1,000,000)
Factual & neutral (43.8%)	4.68	3,128	0.67
Junk & conspiratorial (2.2%)	0.33	2,915	8.89
Personal & investigative (19.1%)	2.46	5,852	2.38
Political (21.3%)	2.42	8,923	3.68

Source: Authors’ calculations based on data collected on 20/03/2020. \*Note: Comments may be disabled on some YouTube videos by the channel owner.

**Figure 1. Distribution of content by channel type**



Source: Authors' calculations based on data collected on 20/03/2020.  
 Note: Figure omits non-applicable and non-English language content

### Public engagement

Average view and comment counts are useful, but imperfect, indicators of video popularity amongst YouTube users. Across the board, videos classified by our team as relaying high-quality factual information accumulate more views than all other content categories, averaging 4.7 million per video, and about fourteen times as many views as junk and conspiratorial videos (see Table 3). Political and personal or investigative videos had both received an average of 2.4 million views at the time of data collection. Politicized content is by far the most commented on, averaging almost 9 thousand comments per clip, compared to just over 3 thousand comments on average for factual content. Conspiratorial and junk content, however, has a higher comment to views ratio than any other category—around 9 thousand comments per million views.

### CONCLUSION

It is important to acknowledge the limitations of this study. Given YouTube's API restrictions and the proprietary nature of its search algorithm and recommender system, we are limited in the kind of claims we are able to make about the representativeness of our sample. Notably, the video sample analyzed in this memo does not account for the personalization and localization of YouTube search rankings. Nevertheless, following the steps outlined

above, we can approximate what an average user would see on YouTube when using these search terms.

People search for health information on YouTube for a number of psychological and social reasons.<sup>[13]</sup> Anxieties surrounding one's own health, current health status, and ability to use the internet, for example, all shape search habits.<sup>[14]</sup> In practice, there is great variety in search skills, internet access, and people search for health related information in different ways. Women, young people, individuals with advanced or college degree, and those living in higher income households tend to pro-actively seek out health news and information online.<sup>[15]</sup>

Finally, it is worth noting that in March 2020 YouTube shifted between demonetizing and monetizing coronavirus-related content for creators under its "sensitive events" policy, which may have caused omissions from our sample.<sup>[16],[17]</sup> It is also possible that some of the prominent creators in our sample extended the reach of their content through YouTube advertising before those restrictions were put in place, which would have improved their performance and discoverability in organic search.

Our study sought to determine the types and informational quality of video content returned by YouTube for different search queries related to coronavirus. We find that: 1) four-fifths of the channels sharing coronavirus news and information are maintained by professional news outlets and that the

channels of public health agencies are rarely, if ever, returned with search results; 2) searches for popular coronavirus-related terms return mostly factual and neutral video results, with low volumes of conspiratorial

or junk science video results; 3) highly politicized health news and information receives on average more public engagement in the form of comments than any other type of videos.

## REFERENCES

- [1] COVID-19 Map', *Johns Hopkins Coronavirus Resource Center*, 2020. <https://coronavirus.jhu.edu/map.html>.
- [2] Institute for Strategic Dialogue, 'COVID-19 Disinformation Briefing No.1', Institute for Strategic Dialogue. [Online]. Available: <https://www.isdglobal.org/isd-publications/covid-19-disinformation-briefing-no-1/>.
- [3] WHO, 'Novel Coronavirus (2019-nCoV) situation reports'. Accessed: Apr. 01, 2020. [Online]. Available: <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports>.
- [4] M. Anderson, and J. Jiang, 'Teens, Social Media & Technology 2018', *Pew Research Center: Internet, Science & Tech*, May 31, 2018. <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>.
- [5] Ofcom, 'Media Nations 2019', Ofcom, 2019. [Online]. Available: <https://www.ofcom.org.uk/research-and-data/tv-radio-and-on-demand/media-nations-2019>.
- [6] A. Przybylski and P. Etchells, 'Don't Freak Out About Quarantine Screen Time', *The New York Times*, Apr. 06, 2020.
- [7] M. Bendersky, L. Garcia-Pueyo, J. Harmsen, V. Josifovski, and D. Lepikhin, 'Up next: retrieval methods for large scale related video suggestion', in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, New York, USA, Aug. 2014, pp. 1769–1778, doi: 10.1145/2623330.2623344.
- [8] S. Bradshaw and P. N. Howard, 'Why Does Junk News Spread So Quickly Across Social Media? — Oxford Internet Institute', Knight Foundation, Miami, 2018. Accessed: Jun. 12, 2019. [Online]. Available: <https://www.oii.ox.ac.uk/blog/why-does-junk-news-spread-so-quickly-across-social-media/>.
- [9] V. N. Gudivada, D. Rao, and J. Paris, 'Understanding Search-Engine Optimization', *Computer*, vol. 48, no. 10, pp. 43–52, Oct. 2015, doi: 10.1109/MC.2015.297.
- [10] A. Rodriguez, 'YouTube's recommendations drive 70% of what we watch', *Quartz*, Jan. 13, 2018.
- [11] 'YouTube Data API Documentation', *Google Developers*, 2020. <https://developers.google.com/youtube/v3/getting-started>.
- [12] X. Cheng, J. Liu, and C. Dale, 'Understanding the Characteristics of Internet Short Video Sharing: A YouTube-Based Measurement Study', *IEEE Trans. Multimed.*, vol. 15, no. 5, pp. 1184–1194, Aug. 2013, doi: 10.1109/TMM.2013.2265531.
- [13] A. Mills and N. Todorova, 'An Integrated Perspective on Factors Influencing Online Health-Information Seeking Behaviours', *ACIS 2016 Proc.*, Jan. 2016, [Online]. Available: <https://aisel.aisnet.org/acis2016/83>.
- [14] C. Lagoe and D. Atkin, 'Health Anxiety in the Digital age: An Exploration of Psychological Determinants of Online Health Information Seeking', *Comput. Hum. Behav.*, vol. 52, pp. 484–491, 2015, doi: 10.1016/j.chb.2015.06.003.
- [15] M. Duggan, 'Health Online 2013', *Pew Research Center: Internet, Science & Tech*, Jan. 15, 2013. <https://www.pewresearch.org/internet/2013/01/15/health-online-2013/> (accessed Mar. 31, 2020).
- [16] S. Perez, 'YouTube will now allow creators to monetize videos about coronavirus and COVID-19', *TechCrunch*, Mar. 11, 2020.
- [17] J. Alexander, 'YouTube is demonetizing videos about coronavirus, and creators are mad', *The Verge*, Mar. 04, 2020.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the European Research Council for the project 'Computational Propaganda: Investigating the Impact of Algorithms and Bots on Political Discourse in Europe', Proposal 648311, 2015–2020, Philip N. Howard, Principal Investigator. Project activities were approved by the University of Oxford's Research Ethics Committee, CUREC OII C1A 15-044. We are also grateful to the Hewlett, Adessium, Luminate, and Ford Foundation for supporting this work. Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the University of Oxford or our funders. We are grateful to Kate Joynes-Burgess, Lisa-Maria Neudert and Dr. Vidya Narayanan for their contributions to this memo.

## ABOUT THE PROJECT

The Computational Propaganda Project (COMPROP), which is based at the Oxford Internet Institute, University of Oxford, involves an interdisciplinary team of social and information scientists researching how political actors manipulate public opinion over social networks. This work includes analyzing how the interaction of algorithms, automation, politics, and social media amplifies or represses political content, disinformation, hate speech, and junk news. Data memos integrate important trends identified during analyses of current events with basic data visualizations, and although they reflect methodological experience and considered analysis, they have not been peer reviewed. Working papers present deeper analysis and extended arguments that have been collegially reviewed and engage with public issues. COMPROP's articles, book chapters, and books are significant manuscripts that have been through peer review and formally published.