YOUTUBE'S SAFETY: LOST IN TRANSLATION

Youtube is not consistently applying safety features across all the languages it supports.

AI FORENSICS

Table of Contents

Credits

Authors: Daniel Jurg, Salvatore Romano, Raziye Buse Çetin, Dr. Bernhard Rieder

Graphic & Brand Design: Denis Constant / Ittaï Studio

Email: info@aiforensics.org

The contribution from AI Forensics is funded by core grants from Open Society Foundations, Luminate, and Limelight Foundation. The contribution also resulted from initial funding by the Dutch Ministry of Internal Affairs. Fonds Wetenschappelijk Onderzoek provides funding for Daniel Jurg as PhD candidate at imec-SMIT, Vrije Universiteit Brussel (grant number FWOM1133).

Executive Summary

Among its various efforts to combat mis- and disinformation, YouTube introduced Topical Context in Information Panels (TCIPs) and Information Panels Providing Publisher Context (IPPPCs). These panels aim to provide context for searches on contentious topics or specific channels types by displaying additional information sourced from independent, third-party partners. In their systemic risk assessments, Google argues that context information panels function as an important feature in maintaining user trust and safety across Google's services, like Search and YouTube, by providing additional background that empowers users to make more informed decisions about the content they encounter and navigate potential harms. This report examines 16 of such context panels across all 83 languages supported by YouTube.

Our findings reveal an uneven distribution of contextual information across languages available on YouTube. This can result in unequal support for users to evaluate content trustworthiness in non-Western regions of the world but also among European citizens voting in European elections or watching government-funded news on YouTube.

Some of the key findings include:

- **Within the European Union's single market, YouTube seems to apply its information panels unevenly.** For example, publicly funded news outlets such as Euronews are labeled for their funding from the European Union in some countries but remain untagged in others, including Denmark, Greece, Iceland, Finland, Norway, and Portugal.

- **The only language displaying the 16 panels** we checked including the recently introduced fact-checker label for fact-checking organizations **is English.** There are only three other languages (Spanish, French and German)

having all the banners except the fact-checker label, and only 2 out of the 16 panels are present for all the languages, showing marked inconsistencies. YouTube's application of information panels suggest disproportionate attention towards Western languages, while neglecting large language communities in parts of Africa, Southeast Asia, and regional or minority languages like Basque, Catalan, and Galician.

- In its previous transparency report under the EU's Code of Practice on Disinformation, YouTube disclosed only the number of impressions of these TCIPs and IPPPCs per EEA country. While this provides a general sense of reach, it is insufficient for assessing how often the panels were missing in instances where they should have been shown, or whether the measure was applied consistently across EEA countries.

- Although YouTube's information panels typically align with a user's language settings, some appear also geographically restricted. For example, panels for abortion-related searches appear only to users within the United States. This suggests that for certain topics, access to this contextual information depends on a user's physical location (or IP address via a VPN) rather than their selected region and language settings.

- YouTube explicitly states that "information panels may not be available in all countries/regions and languages," and that it is "working to bring information panels to more countries/regions." However, labels such as funding information were already introduced in 2018, and there is still limited transparency about the current availability of TCIPs and IPPPCs across the full range of languages and regions in which YouTube operates, as well as about its efforts to expand them.

Given YouTube's stated commitment under the EU's Code of Practice on Disinformation "to assess and update the topics prone to misinformation that receive additional context from information panels," transparency regarding the

availability of these **panels across regional and linguistic contexts seems essential (although not compulsory).** Our aim is to warn lawmakers that such context banners are neither universally applied nor consistently available, to support appropriate steps in the right direction. We developed a dashboard accessible at https://www.ytpanelmonitor.com/. This tool allows users to monitor the availability of context labels across various languages, empowering them to make informed decisions regarding the safety implications for their chosen language.

Disclaimers

This project used automated scrapers to capture the availability of Topical Context in Information Panels (TCIPs) and Information Panels Providing Publisher Context (IPPPCs). While multiple tests were run, the results in this report provide a snapshot at a single moment in time. Therefore, the availability of TCIPs and IPPPCs might have changed since the moment of capture. Since panels are publicly visible and users can adjust their language settings, all findings can be independently verified by entering relevant queries and switching languages. We build a dashboard that regularly checks the availability allowing insights into any future changes.

This report draws on publicly accessible data that we directly observed, but it takes a specific sample of panels we identified ourselves, meaning the findings reflect only a partial view of YouTube's broader practices. This report is not exhaustive and certain elements that were not publicly available and/or not analysed may have been missed. The sole purpose of this report is to explain some of YouTube's practices concerning its safety features without accusing YouTube of not complying with the law or regulations or asserting any violation of law.

Introduction

As part of a study funded by the Dutch Ministry of Internal Affairs on content moderation (see pre-print), we examined how YouTube handled moderation during the 2024 European Parliament Elections. One key finding suggested YouTube's prioritization of legacy media and Public Service Media (PSM) in searches of election-related content on the platform. This drew our attention to the special treatment that PSM channels get on YouTube. PSM channels receive Panels Providing Publisher Context (IPPPCs), which in our initial study seemed to have some inconsistent application. During data collection, we sought to capture which channels displayed so-called funding notices and which did not. Surprisingly, we found that the expected IPPPCs were absent in Portugal, one of the countries where we were accessing YouTube from. Starting from there, similar inconsistencies were observed in the availability of other information panels in different languages. For instance, the panel related to the Covid-19 pandemic was not available to users from The Netherlands navigating the platform in Dutch.

To explore the reason behind this unexpected phenomenon, we tested the availability of panels using VPNs, adjusted location and language settings. These manual tests revealed that it would not be the location (with and without VPN), but the specific language setting that determined whether a feature was available or not. In the case of IPPPCs, panels were missing in several languages including Suomi (Finnish), ελληνικά (Greek), Dansk (Danish), Català (Catalan), Euskara, (Basque), Galego (Galician), and Português (Portuguese). We contacted colleagues in Denmark and Greece, and they confirmed that the panels did not appear and even asked whether this was a "new concept," although we know that the feature was introduced on Youtube in 2018. This discovery led us to pursue a deeper systematic analysis of how Topical Context in Information Panels (TCIPs) and Information Panels Providing Publisher Context (IPPPCs) are made available across all 83 languages supported by YouTube. The findings are presented in this report.

Methodology

Building on previous insights about the IPPPCs being tied to language, we expanded our research to check the availability of TCIPs and IPPPCs systematically across all 83 distinct language settings and subsequently tied to their availability in countries based on language communities. Later on, we tested a sub-sample of 50 countries, modifying not only the language settings, but also the IP location with a VPN.

Query Design

It was not possible to exhaustively analyze all the queries that generated TCIPs as, to our knowledge, there is no point of reference for all searches receiving such panels. Below, we outline the rationale behind our search selection.

Since YouTube references Wikipedia for additional information on specific conspiracy theories, Wikipedia's list of conspiracy theories was used as the basis for the query list. With English (US) as the baseline for panel availability, searches were conducted using the headers from the list. In cases where a header was too abstract, subcategories were also considered (e.g., within 'Antisemitism,' the subcategory 'reptilian humanoid' was included). See Figure 1 as an example of TCIP for querying 'reptilian humanoid.'

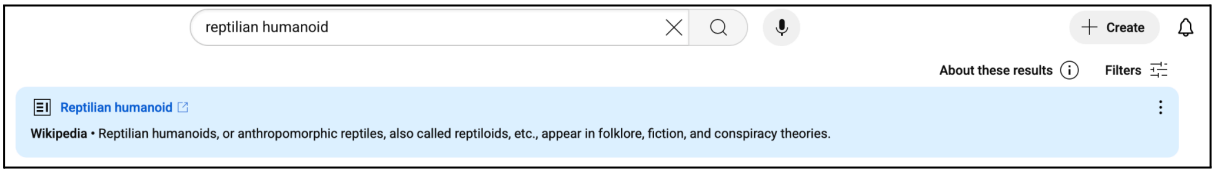

Figure 1. Screenshot for TCIP for searching reptilian humanoid showing TCIP with link to Wikipedia.

Our approach resulted in the selection of twelve queries that generated TCIPs (see Table 1).

Query	Query URL
Chemtrails	www.youtube.com/results?search_query=Chemtrails
Climate Change	www.youtube.com/results?search_query=Climate+Change
COVID-19 pandemic	www.youtube.com/results?search_query=COVID-19+pandemic
Flat earth	www.youtube.com/results?search_query=Flat+earth
Freemasonry	www.youtube.com/results?search_query=Freemasonry
Illuminati	www.youtube.com/results?search_query=Illuminati
John F. Kennedy assassination	www.youtube.com/results?search_query=John+F.+Kennedy+assassination
Malaysia Airlines Flight MH370	www.youtube.com/results?search_query=Malaysia+Airlines+Flight+MH370
New World Order	www.youtube.com/results?search_query=New+World+Order
qanon	www.youtube.com/results?search_query=qanon
September 11 attacks	www.youtube.com/results?search_query=September+11+attacks
Reptilian Humanoid	www.youtube.com/results?search_query=Reptilian+Humanoid

Table 1: Queries used to test TCIPs

While all queries appear directly below the search bar, COVID-19 has more instructive information from the World Health Organization that appears when one scrolls down a couple of videos.

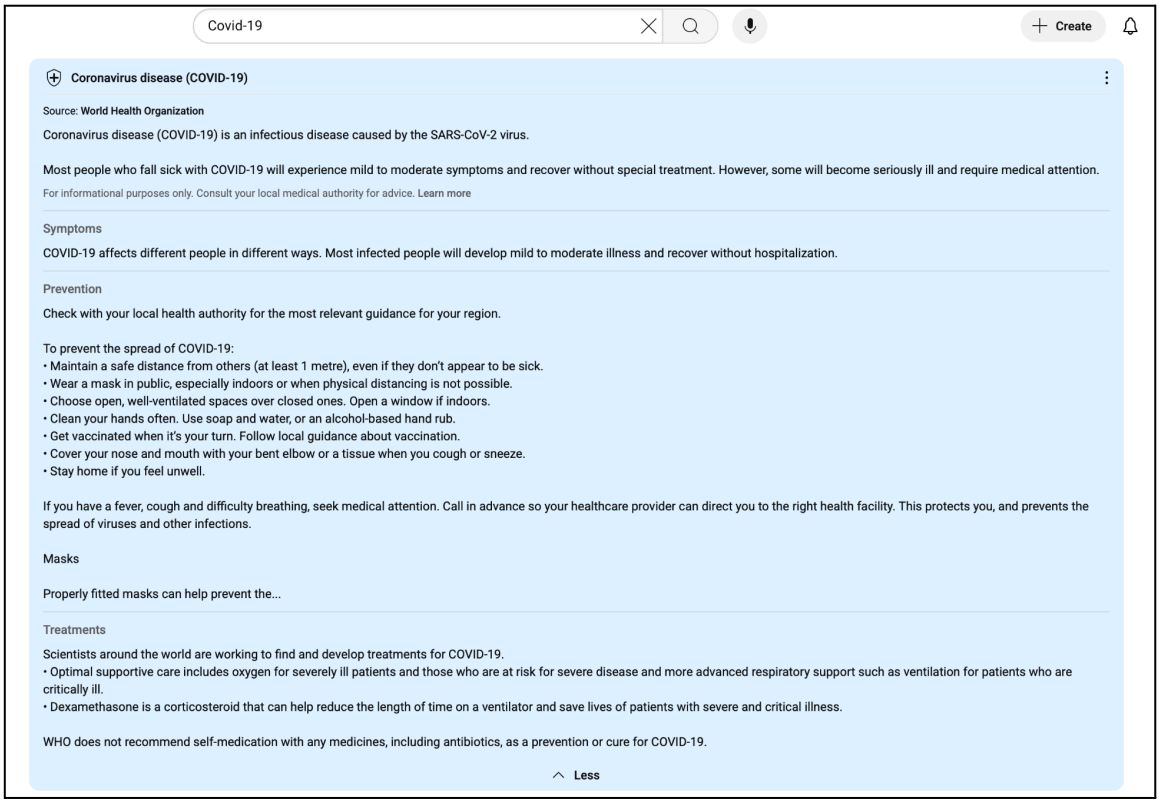

Figure 2. Screenshot for TCIP for searching Covid-19 with information from the World Health Organization.

Finally, in order to find videos with IPPPCs we looked at YouTube's own communication and found information for the following labels being applied on the channel level: (1) Licensed doctors, (2) National health authorities, (3) News funding, and (4) Fact-checkers.

Topic	Channel	URL
News Funding	EuroNews	https://www.youtube.com/watch?v=robbdct9y70
Fact Checker	Fact Check	https://www.youtube.com/watch?v=4Vek73X1jMY
National Health Authority	UK Health Security Agency	https://www.youtube.com/watch?v=SMgM8SoivbU&t=5s&ab_channel=UKHealthSecurityAgency
Licensed Doctor	Healthcare Triage	https://www.youtube.com/watch?v=EUaemzMYWeg

Table 2: Videos used to test IPPPCs

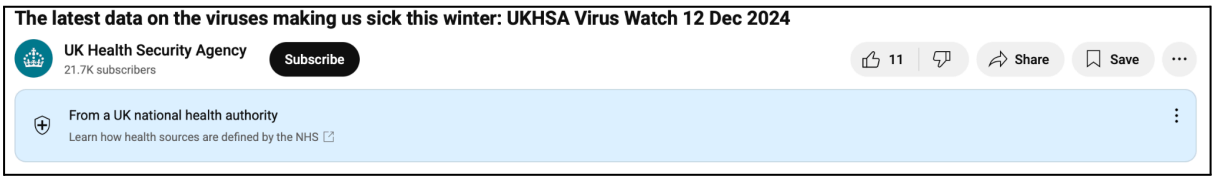

Figure 3. Screenshot for IPPPCs for video from the UK Health Security Agency.

Data Collection

In order to automatically test the availability of TCIPs and IPPPCs, we built a custom scraper using Playwright. We first collected a list of all available languages using YouTube's API, which provided us with 83 distinct languages. We also mapped the 110 regions where YouTube operates in order to link available languages to specific regional contexts. Our custom-built scraper accepts either a query (for TCIPs) or a video ID (for IPPPCs), along with a CSS selector that identifies the panel element. It then iterates over all 83 language settings, performing a YouTube search in each by setting the appropriate language via cookie settings. For each search, the scraper checks for the presence of the specified panel using the CSS selector and records the result as either 'True' or 'False'. The final dataset includes results for 12 queries and 4 videos across all 83 languages (total n=1 328 requests). To ensure reliability, the scraper included detailed logging to flag issues with selector detection, and screenshots were captured from each search to visually verify the output.

Data Analysis

Our initial analysis employs a straightforward approach, utilizing a heatmap-like visualization to compare the availability of TCIPs and IPPPCs across all languages. To further examine the availability and regional discrepancies, a world map visualization was generated using the Plotly library.

Publicly available country information data was first obtained from GeoNames, containing country names and the languages spoken in each country. This data was then used to map TCIPs and IPPPCs availability based on the languages spoken in

each country in which YouTube operates. Countries were classified into three categories:

- **"Available"** if the specific panel was present in all spoken languages,
- **"Partially Available"** if the specific panel was present for some but not all languages, and
- **"Unavailable"** if no spoken language in the country had panel availability.

While this classification provides a perspective on regional variations in TCIPs and IPPPCs availability, there is an inherent limitation when mapping languages to countries. GeoNames, for example, has the following languages listed for France:

- **French** (*fr-FR*)
- **Franco-Provençal (Arpitan)** (*frp*)
- **Breton** (*br*)
- **Corsican** (*co*)
- **Catalan** (*ca*)
- **Basque** (*eu*)
- **Occitan** (*oc*)

In our mapping, we only took into account the 83 languages that are available on YouTube. In the case of France, this is French (fr), Catalan (ca), and Basque (eu). This implies that if a TCIP or IPPPC is available in French, but not in Catalan or Basque it is designated as "partially available" for France. YouTube also uses 'es-149,' which is a code for Spanish language spoken in Latin America and the Caribbean. GeoNames does not use this code. For instance, for Mexico (MX), GeoNames uses 'es-MX.' We have, therefore, sourced a list of Latin American and Caribbean countries that have been part of 'es-419' to map this language. Mapping languages onto countries comes with various trade-offs and issues, and our main aim here is to give a more helpful visual impression of the implications of different degrees of availability in the world.

Limitations

As there is no known exhaustive overview of TCIPs and IPPPCs, it is possible that additional panels exist beyond those we examined. Moreover, YouTube may adjust the availability of these panels over time, meaning the results captured on April 14 (queries) and April 16 (videos) may not reflect the situation after this date and the release of this report. We ran our query analysis on the 3rd of January and on the 14th of April 2025 and found some changes in relation to COVID-19 and Climate change availability. In most cases, we did not find the same COVID-19 banner at the top suggesting that YouTube has moved away from prioritizing information on the topic. However, the online video platform does still display the banner from the World Health Organization that appeared after the first couple of videos, which we used to test COVID-19 context information. Moreover, re-running the analysis shows that YouTube is expanding the availability of Climate change information, for example, in French, French (Canadian), Korean, Malay, Slovenian, Turkish, Vietnamese. Our dashboard is designed to monitor changes in YouTube's implementation and availability of TCIPs and IPPPCs: https://www.ytpanelmonitor.com/. Our dashboard also incorporates additional panels identified after the initial analysis. It is important to notice that our data is captured on a virtual machine in Belgium without using a VPN. We acknowledge this setup could yield results that differ from those in other regions, like the United States. To account for this, we compared our findings against a 50-language sample using different VPNs, which revealed only minor discrepancies. A comprehensive test of all 83 languages across all 110 regions using different VPN settings, however, remains beyond the scope of this project. Our main goal here is to note the observed inequalities in the application of context panels.

Finally, this report only concerns the availability of specific information panels, but questions concerning what kinds of topics get information panels is something that we do not engage with. We are also not aware of specific information banners that are available in other languages and not in English. Nonetheless, the primary aim of this research is to highlight the (un)availability of TCIPs and IPPPCs and the significant disparities across languages and regions.

Findings

This section presents the results of our TCIP and IPPPC tests across all YouTube interface languages, based on twelve queries and four videos. We categorize the outcomes according to most covered languages, mixed coverage, and excluded languages. Our dashboard offers an interactive environment in which the availability of all individual panels can be explored with updated information.

An initial observation about context panel availability is that no other language receives the full set of panels provided to US English users. This is due to YouTube's recent introduction of a fact-checker label for verified fact-checking channels, a feature not yet extended to other language markets. Two video-based labels are applied universally: *National Health Authority* and *Licensed Doctor*. These labels appear for channels such as the UK Health Security Agency, ensuring that users are informed when content comes from official health professionals or institutions, regardless of their language settings. This consistency contrasts with the uneven application of other panels. We will now move towards more specific observations, clustering countries on the basis of availability as reflected in the query-based tests for TCIPs.

Most covered languages

Figure 5 highlights nine languages (including national variants) that receive comparable treatment to US English users.

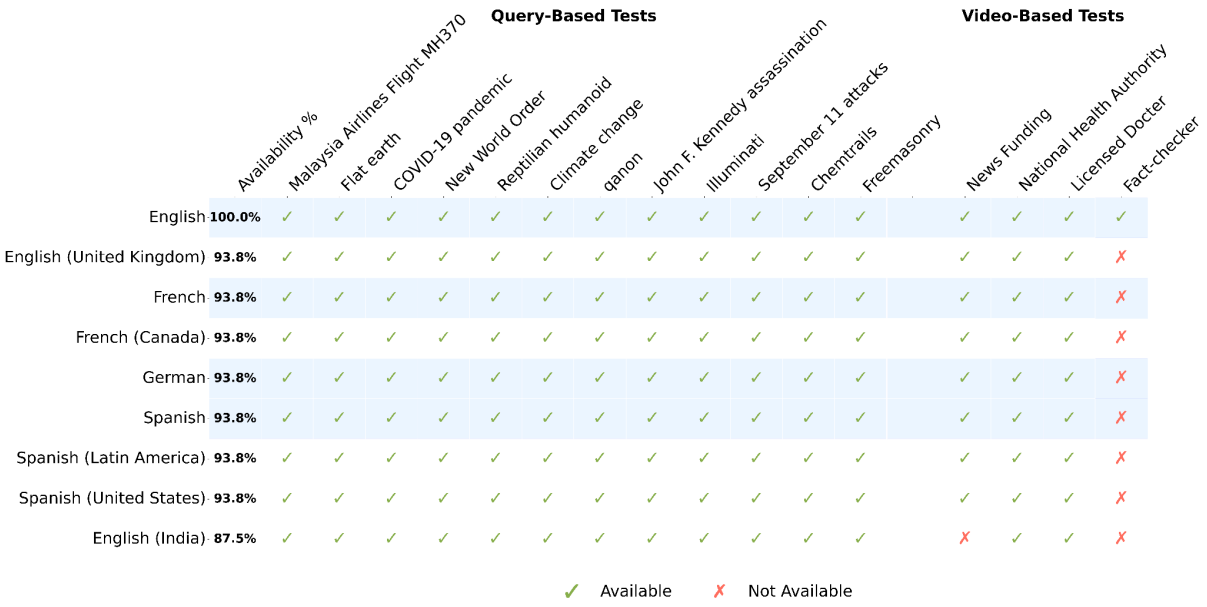

Figure 5. Most covered languages.

Figure 5 reveals a strong availability for major European languages (highlighted in blue) such as French, German, and Spanish. It also accounts for the latin-speaking user-base and Indian users who have their settings to Indian English. All of these language communities receive the same features as US English users except that Indian English users do not receive information about the funding of news organizations such as public service media or specific government funded channels. This means, for example, that users in all countries in the most covered languages can see that Doordarshan is India's state-owned public television broadcaster, but Indian English users do not get this information on their YouTube page.

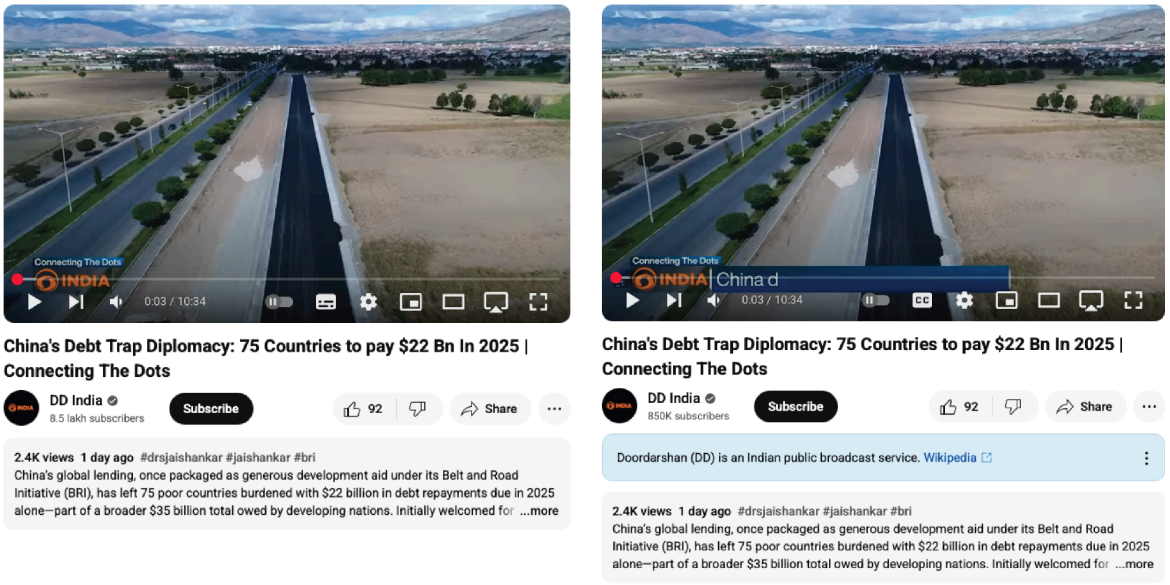

Figure 6. Comparison between English (Indian) view of video from Doordarshan's YouTube channel (left) and English (US) users (right).

Doordarshan primarily operates in Hindi, which does show the banner. However, it also operates in other Indian languages in which the banner is also not visible. While YouTube has made claims that it simply provides information to users about funding, we argue in our content moderation chapter that the News Funding Notice (NFN) arrived within a particular context concerning Russian disinformation threatening US election integrity. Not labelling public service media in all the language(s) of the country in which the organization is based is surprising, given that US citizens and dominant language users are informed of this information. It raises question which users such panels are meant to serve.

Mixed covered languages

We can extend the observation about the NFN's to Europe as well, where we can see that for languages with mixed coverage, even within the European Union's single market (EU and EEA) (highlighted in blue), YouTube applies information panels unevenly, as they are missing in Denmark, Greece, Iceland, Finland, Norway, and Portugal (Fig. 6).

	Availability %	Query-Based Tests: Malaysia Airlines Flight MH370	Flat earth	COVID-19 pandemic	New World Order	Reptilian humanoid	Climate change	qanon	John F. Kennedy assassination	Illuminati	September 11 attacks	Chemtrails	Freemasonry	Video-Based Tests: News Funding	National Health Authority	Licensed Docter	Fact-checker
Korean	87.5%	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Portuguese	87.5%	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Italian	81.2%	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✗
Russian	81.2%	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Indonesian	81.2%	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Chinese (Hong Kong)	81.2%	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Chinese (Taiwan)	81.2%	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Czech	81.2%	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Polish	75.0%	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Greek	75.0%	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗
Arabic	75.0%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Swedish	75.0%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Turkish	75.0%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Bulgarian	75.0%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗
Japanese	68.8%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗
Romanian	68.8%	✓	✓	✗	✗	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗
Slovak	68.8%	✓	✗	✗	✓	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Norwegian	68.8%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗
Ukrainian	68.8%	✗	✓	✗	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗
Serbian	68.8%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗
Vietnamese	68.8%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗
Hebrew	68.8%	✓	✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗
Dutch	68.8%	✓	✓	✗	✓	✓	✓	✗	✗	✓	✓	✗	✓	✓	✓	✓	✗
Thai	62.5%	✓	✗	✓	✗	✗	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗
Hungarian	62.5%	✓	✗	✗	✗	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗
Finnish	56.2%	✓	✗	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓	✗	✓	✓	✗
Croatian	56.2%	✗	✓	✗	✗	✗	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗
Latvian	56.2%	✓	✓	✗	✗	✗	✓	✗	✓	✓	✗	✗	✓	✓	✓	✓	✗
Danish	56.2%	✓	✗	✗	✗	✗	✓	✓	✓	✓	✓	✗	✓	✗	✓	✓	✗
Sinhala	50.0%	✓	✓	✗	✗	✗	✓	✗	✓	✓	✗	✗	✓	✗	✓	✓	✗
Malay	50.0%	✓	✓	✗	✗	✗	✓	✗	✓	✓	✗	✗	✓	✗	✓	✓	✗
Slovenian	50.0%	✗	✗	✗	✗	✗	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✗
Estonian	50.0%	✓	✓	✗	✗	✗	✓	✗	✗	✗	✗	✓	✓	✓	✓	✓	✗
Tamil	43.8%	✓	✓	✗	✗	✗	✓	✗	✗	✓	✗	✗	✓	✗	✓	✓	✗
Bangla	43.8%	✓	✓	✗	✗	✗	✓	✗	✗	✓	✗	✗	✓	✗	✓	✓	✗
Lithuanian	37.5%	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✓	✓	✓	✓	✗
Bosnian	37.5%	✗	✗	✗	✗	✗	✓	✗	✓	✓	✗	✗	✓	✗	✓	✓	✗
Georgian	31.2%	✗	✗	✗	✗	✗	✓	✗	✗	✓	✗	✗	✓	✗	✓	✓	✗
Hindi	25.0%	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗
Nepali	25.0%	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✓	✓	✗
Kazakh	25.0%	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗	✓	✗	✓	✓	✗
Icelandic	18.8%	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✓	✓	✗
Lao	18.8%	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✓	✗

✓ Available ✗ Not Available

Figure 6. Languages with mixed coverage.

What stands out most when examining mixed covered languages is the lack of a consistent pattern in TCIP availability across languages spoken within the same geographical regions. Alongside the NFNs, countries like the Netherlands and Denmark exhibit partial overlap for some panels but diverge in others. This seemingly erratic distribution makes it difficult to draw any conclusions about the intent behind the distribution of context panels. Beyond discrepancies between national languages, we also observe uneven treatment of local (minority) language communities. In Spain, for example, users with their language set to Spanish receive the full set of TCIPs, while those using regional languages like Catalan, Basque, or Galician do not see any banners (see Figure 7).

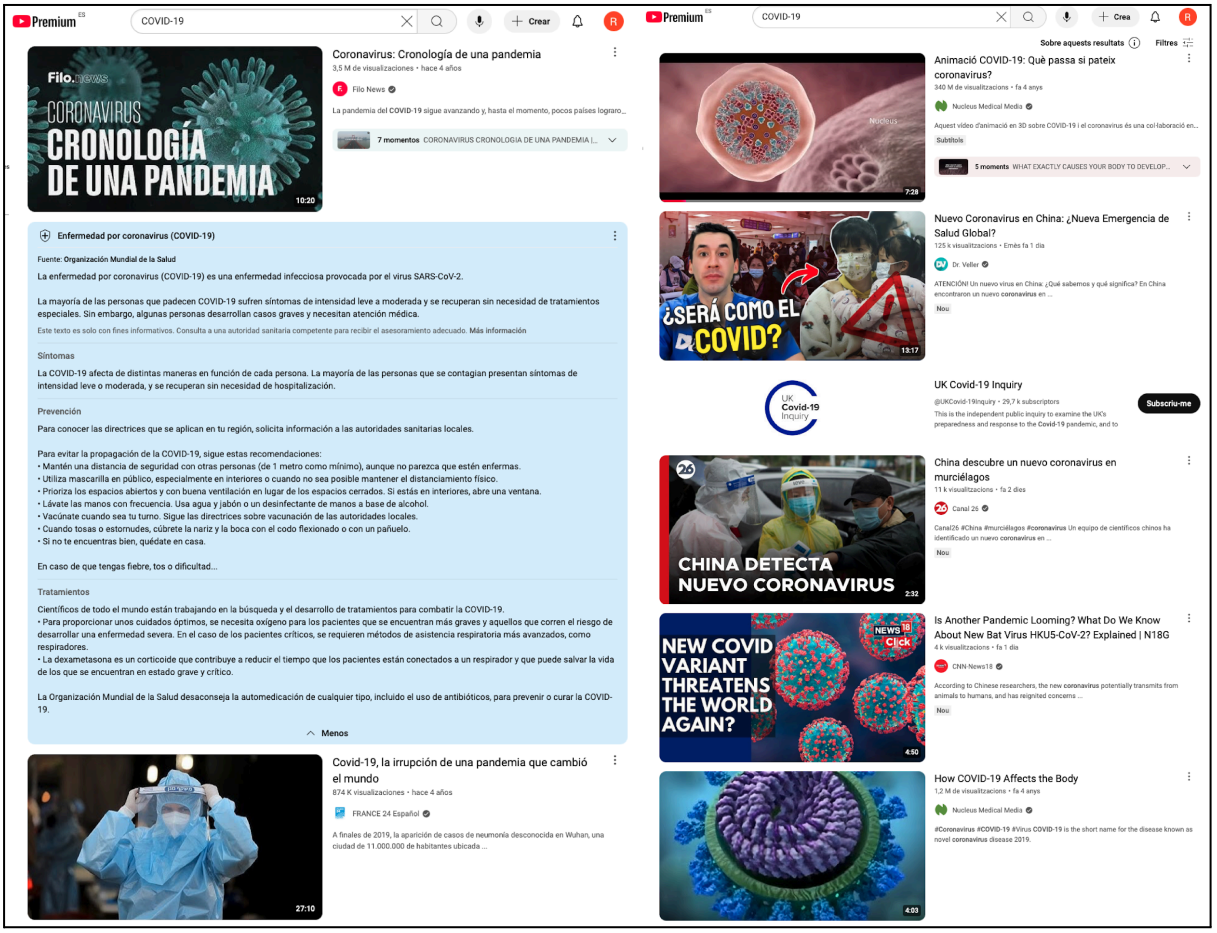

Figure 7: Screenshots of searches for COVID-19 (26-02-2025) contrasting Spanish view (left) with the Catalan view (right).

The figure above seeks to demonstrate that Spanish users who are searching for COVID information will receive detailed information from the WHO about what the symptoms are, how to best prevent the spread, and what the treatments are. In this case, one can argue that a simple translation of the text in a language that is supported individually in a country has important implications for national health concerns.

Excluded languages

Finally, we identify a group of languages that receive no coverage for any of the tested queries, apart from the universally applied labels on official health-related videos (Fig. 8). This group primarily consists of languages spoken in African and South Asian countries, including various regional language communities within India.

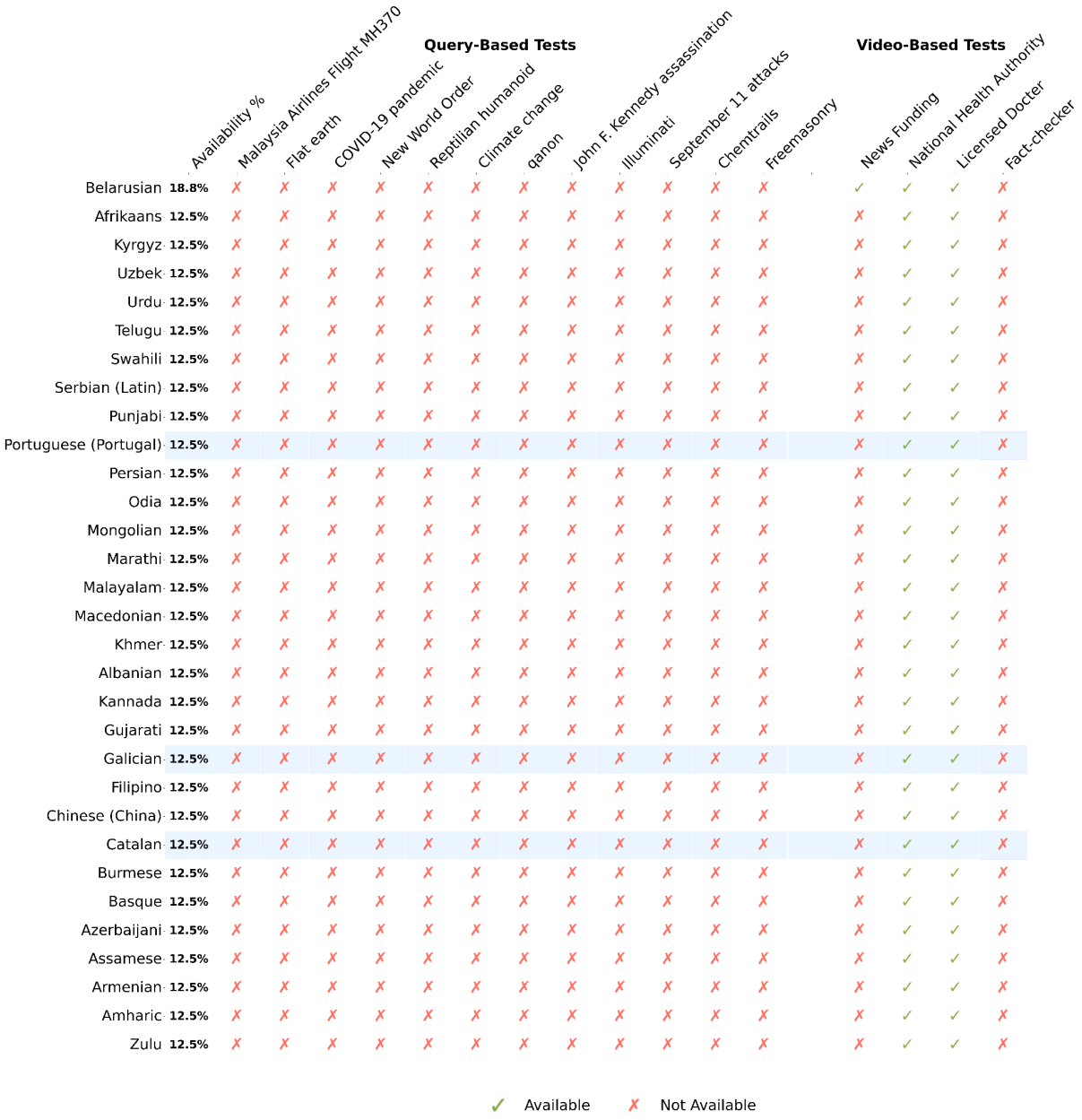

Figure 8. Languages without any context banners for queries.

In order to get some sense of the geographical impact of limiting information, we present Figure 9 for the availability of languages in the markets in which YouTube operates.

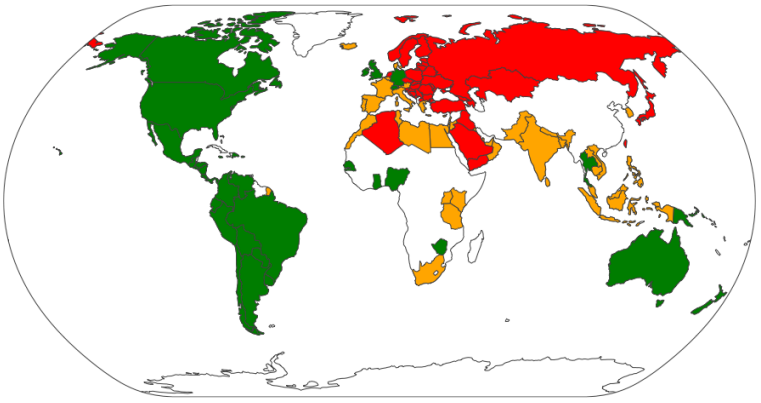

Figure 9*:* Map with countries in which YouTube has regional settings, showing the availability of COVID-19 TCIPs in the languages spoken in that country. Green is full availability, yellow is partial availability, red is no availability and white implies that YouTube does not have a regional setting for that specific country (taken from latest run on our dashboard captured on 2025-07-07). See the interactive version of this graph here: https://ytpanelmonitor.com/

Figure 9 highlights the divide between East and West to access COVID-19 information, showing that Nordic and Eastern European countries do not receive the same information as their Western European counterparts. In addition, it shows how Russia, Kazakhstan and Turkey were not getting the notification, whereas North, Central, and South America were well covered.

Location-based Restrictions

While working on this report, we discovered only near the end that YouTube had a banner for searches related to "abortion" previously undetected by us. We made this observations as we were running the data collection from the United States during a temporary research stay.

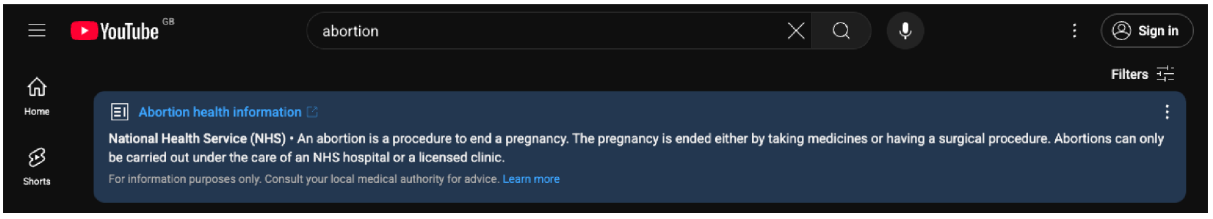

Figure 10*:* Screenshot of abortion-query performed for English on 25-05-03.

This observation revealed that some information panels, such as for searches around abortion, are geo-restricted to specific locations. This finding contrasted with our earlier VPN tests within Europe for the government report, which had not shown such restrictions. Ultimately, such additional location-based restrictions make it even more difficult to determine when, where, and to whom YouTube displays its information panels.

In order to get some insight into these restrictions we aimed to map the differences between languages settings and language settings with additional VPN settings. Given the computational difficulties of capturing panels with VPN settings on a large scale, we were only able to compare a smaller language sample. However, even this small sample did allow us to uncover additional instances where a user's geographic location paradoxically revoked access to localized information panels. Table 3 presents the results of VPN tests where we were able to capture results without bot detection and where the language could be mapped with the available VPN settings.

Query/Video	Language	Country (VPN)	Panel presence with local VPN	Panel presence without local VPN
COVID-19 pandemic	German	Germany	False	True
Qanon	Spanish	Spain	False	True
Armenian genocide	Spanish	Spain	False	True
Armenian genocide	Turkish	Turkey	False	True
September 11 attacks	Danish	Denmark	False	True
September 11 attacks	French	France	False	True
September 11 attacks	Norwegian	Norway	False	True
September 11 attacks	Polish	Poland	False	True

September 11 attacks	Swedish	Sweden	False	True
September 11 attacks	Ukrainian	Ukraine	False	True

Table 3: Discrepancies in panel availability based on geolocation. The fourth column indicates the presence of a banner based on the language and the IP country, while the fifth column shows banner availability with a Belgian IP, maintaining the same language.

Table 3 shows that while a context panel for "9/11" searches is, for example, displayed when the language is Danish and we use a Belgian IP, the panel is suppressed when the user's language is still Danish but the location is moved to Denmark via a VPN. Such logistical inconsistencies mean that even when translations are available, it may still be rendered inaccessible to the target demographic. Another interesting difference is that the Armenian genocide banner[1] is not present if accessed in Turkish while in Turkey, but Turkish language banners of the Armenian genocide can be accessed outside of Turkey.

Our VPN tests only covered the official languages of a subset of countries, using their respective IP addresses. A further test could determine if banners are displayed consistently across countries across all the available languages. This would involve testing all 83 possible languages for each of the 16 banners across approximately 110 countries in which YouTube operates, resulting in 146,080 potential combinations. While this level of complexity is beyond the scope of this report, it is the only way to identify all possible inconsistencies.

Despite such computational limitations, this methodology reveals interesting contradictions and limitations. We discovered a compelling example using such a methodology: the government-funded news banner appears on the Greek YouTube channel @ert when accessed from the US with the language set to Greek. Previously, we observed that using Greek as the language alone did not trigger the banner, and the same was true when using a Greek IP address. Further testing revealed that the banner also does not appear when using English as the language

[1] Note that the Armenian genocide banner was not tested in the previous experiment without VPNs, but only in this experiment. We aim to update our dashboard with new banners as soon as we discover they exist.

with a Greek IP. The ratio behind these inconsistencies is hard to explain, and seems to lead to major problems for users, especially across EU countries. The fact that the Public broadcasters like Euronews display funding labels in some EU countries but remain unlabeled in others, including when using an IP located in Denmark, Greece, Finland, Norway, and Portugal, could be quite problematic.

Assessing the TCIPs and IPPPCs against the EU Digital Standards

The uneven implementation of information panels across languages that our research has identified could raise important questions about YouTube's compliance with existing EU regulatory frameworks designed to ensure equal access to information and protection from harmful content.

Legal and Co-Regulatory Frameworks

YouTube's operations in Europe are governed by an evolving ecosystem of directives, regulations and codes that collectively establish standards for platform responsibility. Rather than functioning as isolated requirements, these frameworks form an interconnected set of obligations centered on user protection and information integrity.

The Audiovisual Media Services Directive (AVMSD) requires video-sharing platforms like YouTube to put in place "appropriate measures" to protect users from harmful content, including disinformation. Such measures include "providing effective media literacy measures and tools and raising users' awareness of those measures and

tools." While the directive does not explicitly mandate topical information panels, the significant disparities in panel implementation across languages we've documented suggest an insufficient fulfillment of these core protective obligations. The directive is transposed into national laws across member states.

This baseline framework is reinforced by the Strengthened Code of Practice on Disinformation, which has now been converted into a Code under the Digital Services Act (DSA) framework. The Code of Practice is developed by the signatories with the guidance of the European Commission and is an example of co-regulation. Its recent conversion into a Code of Conduct under the DSA transforms what were previously seen as voluntary industry commitments into elements with legal significance. While critics have questioned the effectiveness of co-regulatory approaches like Codes of Practice, the DSA creates a new dynamic:

Failure to comply with commitments under the Code of Practice can now serve as an indicator of non-compliance with the risk mitigation obligations in Article 35 of the DSA. Conversely, adherence to these commitments can help platforms demonstrate compliance with these same obligations.

YouTube stated that the TCIPs were part of their efforts under Commitment 17 and 22.7 of the Code on user empowerment and media literacy, as well as the systemic risk mitigation required by the DSA. In a previous report, YouTube only disclosed the number of impressions of these information panels per EEA country. This single metric is insufficient to determine the number of impressions of videos that should have contained the TCPPs but were missed, or whether the measure is consistently applied across countries.

Even though Youtube does not appear to have violated any applicable regulations, it seems to us that Youtube could promote better practices given the importance of these issues.

Conclusion

This study reveals significant discrepancies in YouTube's implementation of Topical Context in Information Panels (TCIPs) and Information Panels Providing Publisher Context (IPPPCs). YouTube itself recognises this fact on their website (see links about the specific features). This could demonstrate that YouTube has a policy of selective implementation of disinformation features to enhance contextual awareness. Whereas many labels are not universally available across all languages and regions, YouTube does consistently label channels belonging to national health authorities or licensed doctors across all languages. The fact that YouTube universally applies labels about the official credentials of certain channels while its implementation of information on conspiratorial topics, and other politically sensitive topics such as abortion, is inconsistent, seems to suggest the online video's platform uneasiness with moderating such issues consistently across their services.

Recommendations

We believe that YouTube could improve its practices in the following ways and encourage lawmakers to legislate accordingly :

- YouTube should demonstrate compliance and efficiency of the TICPs through more meaningful metrics rather than the number of impressions of videos with TICPs per country. This could, for instance, include disclosing financial and human resources invested in the implementation of this measure by country.
- The transparency and risk assessment reports should establish a clear timeline and targets in addressing disparities until next reporting phases.
- Lastly, Youtube should implement equal protections for speakers of regional and minority languages.

In general, researchers, policymakers, and civil society must recognize that these features are neither static nor universally applied. It raises again the need for increased transparency and accountability in platform governance.