

REPORT • OCTOBER 2026

Inconsistencies in climate dis/misinformation moderation on YouTube and TikTok

A phased analysis on video-based platform enforcement of climate disinformation under the EU's Digital Services Act

BY FUNDACIÓN MALDITA.ES AND SCIENCE FEEDBACK
WITH THE SUPPORT OF THE EUROPEAN FACT-CHECKING STANDARDS NETWORK

Index

Executive Summary.....	2
1. Introduction.....	3
Impact and Risks of Climate Dis/misinformation.....	4
TikTok and YouTube: Policy Gaps and Enforcement Failures.....	4
Policy Implications.....	5
2. Overall Findings.....	6
Artificial weather manipulation and chemtrails are the most common climate narrative.....	6
Phase 1: Enforcement of platform policies.....	7
Phase 2: Account-level enforcement.....	8
Phase 3: Appeal mechanism.....	9
3. Findings by Language and Platform.....	10
4. Recommendations.....	14
Recommendations for Platforms.....	14
Recommendations for Regulators.....	15
Recommendations on Out-of-Court Dispute Settlement Bodies (Art. 21 DSA).....	16
5. Methodology.....	17
ANNEX I.....	21
ANNEX II.....	23

Executive Summary

We reviewed 331 videos on TikTok and 894 on YouTube in Spanish and in French, retrieved through a keyword search of terms recurrent in Science Feedback's and Fundación Maldita's climate dis/misinformation work and AI-filtered to flag only those likely to violate platforms' policies. After manually reviewing and fact-checking them, we reported 87 videos to TikTok and 82 to YouTube across both languages.

Key Findings:

- **YouTube does not enforce its policy of demonetising climate dis/misinformation:** after we reported 82 videos to YouTube that contained climate misinformation in violation of the platform's policy, the platform only demonetized less than 5% of those videos.
- **YouTube's climate dis/misinformation moderation is completely opaque:** even if certain videos were seemingly demonetised following our reports, neither Fundación Maldita.es nor Science Feedback received any response.
- **TikTok's moderation of climate dis/misinformation is inconsistent and likely automated:** of the 87 reported videos, TikTok restricted only 20%, focusing on climate-related conspiracy theories. A 0% rate of reversed decisions after appeal strongly indicates purely automated means.
- **Appealing to an ODS and a DSC showed very limited success:** for every negative decision by YouTube and TikTok, we filed a complaint in Spain to a DSA Out-Of-Court Dispute-Settlement Body (ODS) and in France to the national Digital Services Coordinator (DSC). The ODS replied to just 10% of complaints in less than a month, while the DSC did not offer a reply. In France, at the time of writing, the DSC (Arcom) had not replied to the two complaints (one per platform). We will update this report if and when we hear back.
- **Our recommendations** include applying existing policies diligently and consistently, as DSA requires, but also updating them to provide clarity around more sophisticated "new denial" narratives that go beyond the traditional outright denialism. Also: more transparency and human review.

Phased testing framework

Our investigation examines both content-level enforcement (demonetisation on YouTube and removal on TikTok) and account-level enforcement (e.g. suspensions for repeated violations), as well as the functioning of DSA complaint-handling and appeal mechanisms. To do so, the analysis was conducted in **3 phases**.

1

Enforcement of platform policies:

Test how YouTube's advertiser rules and demonetisation system, and TikTok's misinformation and content-removal policies, are enforced in response to user-generated reports of climate disinformation on high-impact accounts.

2

Account-level enforcement:

Assess how platforms respond to repeated user-generated reports of climate disinformation posted by the same accounts over time, and compare content-level and account-level enforcement.

3

Appeal mechanism:

Test DSA out-of-platform appeal mechanisms for climate dis/misinformation cases, in France via article 53 noncompliance complaints directed to the Digital Services Coordinator (DSC) Arcom, and in Spain through complaint submissions to the Out-of-court Dispute Settlement (ODS) [Appeals Centre Europe](#).

1. Introduction

The headline of a YouTube video with over 140,000 views reads “The Climate Consensus Fraud”. It features a series of interviews with Spanish-speaking ‘scientists’ where they dispute the scientific consensus on climate change and claim that anthropogenic CO₂ is not the main cause of global warming, calling the scientific consensus a form of censorship and even going so far as to assert CO₂ is harmless and beneficial. In another video that accumulates more than half a million views, this time on TikTok and in French, a medium abandons scientific explanations altogether and instead claims climate change results from the use of weather manipulation technologies controlled by the government along with an “ancient karmic cycle”.

These are just examples of a larger issue. Politically charged arguments and institutional distrust have influenced the evolution of climate dis/misinformation beyond the outright denial that we have illustrated above. Climate dis/misinformation has grown into a much more complex set of narratives, including ones that may acknowledge aspects of the scientific consensus while questioning the severity or impacts of climate change, dismissing the effectiveness of proposed solutions, or promoting conspiracy theories about the actors and institutions involved in climate action. This has been called the [“new denial”](#), and these types of claims are increasingly circulating on video platforms. In this study, we test whether YouTube and TikTok do take action when a user reports videos that contravene their narrowly written policies on climate dis/misinformation.

Impact and Risks of Climate Dis/misinformation

Climate dis/misinformation has been shown to cause serious harm. The Intergovernmental Panel on Climate Change (IPCC) has [long acknowledged](#) that dis/misinformation in this realm can undermine science and lead to “misperceptions of the scientific consensus, uncertainty, disregarded risk and urgency, and dissent”. Similarly, the United Nations (UN) has also [weighed in on the dangers of dis/misinformation](#), noting that misleading narratives are delaying urgent action against climate change. This delay carries consequences both immediate and lasting. In the near term, climate change can endanger those affected by [extreme weather](#), brought on by or made worse by climate change, by [encumbering the work of government and first responders](#). In the longer term, it can [undermine political will and public confidence](#) to take the critical steps required to avert the worst effects of climate change.

Digital platforms are not oblivious to these risks or the presence of climate dis/misinformation on their services and have respectively voiced their commitment to reducing the spread of this content online. In 2021, YouTube [announced](#) measures to prohibit ads for content that denied the existence or causes of climate change. In their words, this was an initiative among many others that the company has undertaken over decades to “promote sustainability and actively combat climate change”. Similarly, on the occasion of COP28 in 2023, TikTok [recognised](#) climate change as “one of the most important challenges facing humankind” and committed to “raising sustainability awareness and increasing climate literacy” among TikTok’s global community.

Although this enthusiasm seems to have waned on the part of platforms according to [several studies](#), [including ours](#), the volume of climate dis/misinformation content apparently has not. In fact, climate dis/misinformation has caught on like fire on social media, aided in large part by [social media](#) algorithms that reward incendiary content and platforms’ advertising infrastructure that funds disinformation spreaders. There are [numerous studies](#) out there that reveal how consistently platforms [prioritise, amplify and incentivise climate dis/misinformation](#).

TikTok and YouTube: Policy Gaps and Enforcement Failures

TikTok is one of the only platforms that sets out a [content policy against climate dis/misinformation](#). It prohibits “misinformation that denies the existence of climate change, misrepresents its causes, or contradicts its established environmental impact”. However, the “new denial” content, conspiracy theories in particular, do not necessarily fall under their climate dis/misinformation policy. Instead, it falls under adjacent policies, governing conspiracy theories more broadly, that carry lower penalties than the climate dis/misinformation policy, which is directly not allowed on the platform. YouTube, on the other hand, doesn't have content rules

against climate dis/misinformation, but it claims it doesn't allow creators to monetise content containing climate dis/misinformation under its ad policy. Even with these policies in place, both platforms are failing to enforce them, as we have documented in this study.

We focus on YouTube and TikTok specifically because these platforms are widely used in Europe and can play an important role in shaping public perceptions of climate change, particularly among young audiences. At the same time, they have received comparatively less scrutiny than [text-based platforms](#) from fact-checkers and researchers regarding misinformation, partly because monitoring and processing video content is time-consuming. The fact that these platforms rely on algorithmic promotion and that creators often monetise their videos makes them particularly relevant for assessing systemic risk mitigation under the DSA.

Policy Implications

Under the Digital Services Act (DSA), Very Large Online Platforms (VLOPs) such as TikTok and YouTube have the obligation to identify, assess, and mitigate systemic risks arising from the functioning and design of their services. Disinformation is explicitly listed among those risks, and it intersects with other systemic risks like public security. Take, for instance, videos claiming that CO₂ is harmless or false narratives that spread in the aftermath of natural disasters.

While TikTok's latest [DSA risk assessment](#) points to climate dis/misinformation as a public security risk and lays out its collaboration with independent fact-checking partners and climate experts, it fails to recognise how [algorithms on their platform push this type of content](#) and whether its monetisation program rewards creating it. YouTube's [systemic risk assessment](#) simply fails to mention climate dis/misinformation altogether.

It's precisely the gap between what they claim to do, both in their policies and risk assessment, that this research sets out to test. It examines the application of terms and conditions, complaint handling, out-of-court dispute settlement, and risk mitigation. It also assesses self-regulatory adherence to the EU Code of Conduct on Disinformation, of which YouTube and TikTok are signatories, including the commitments to reduce the amplification of misleading content and demonetise dis/misinformation. As such, findings can inform regulators on monetisation of dis/misinformation, effectiveness of enforcement mechanisms, and platform accountability overall.

2. Overall Findings

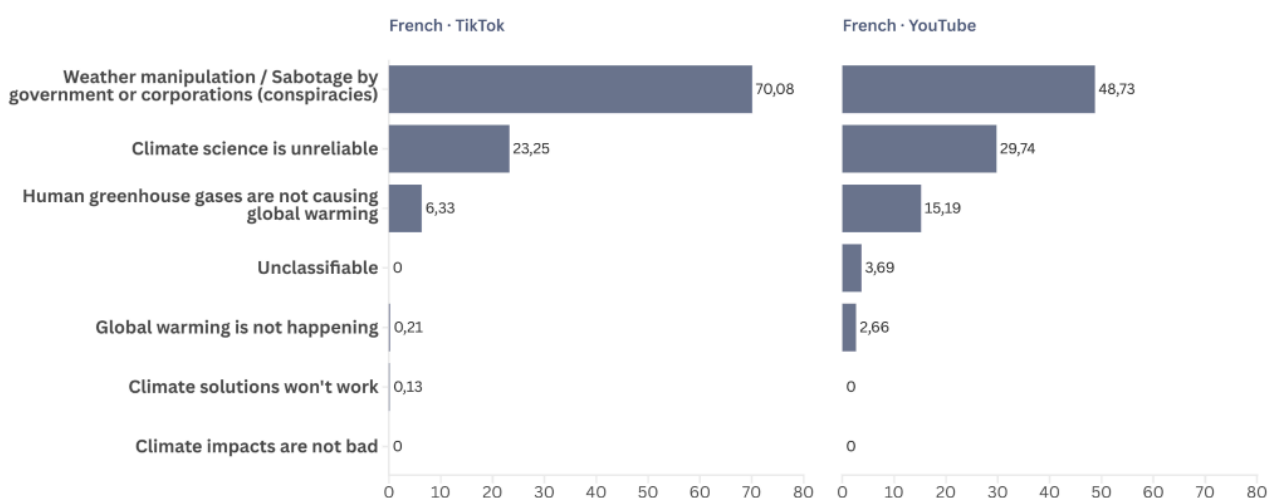
Artificial weather manipulation and chemtrails are the most common climate narrative

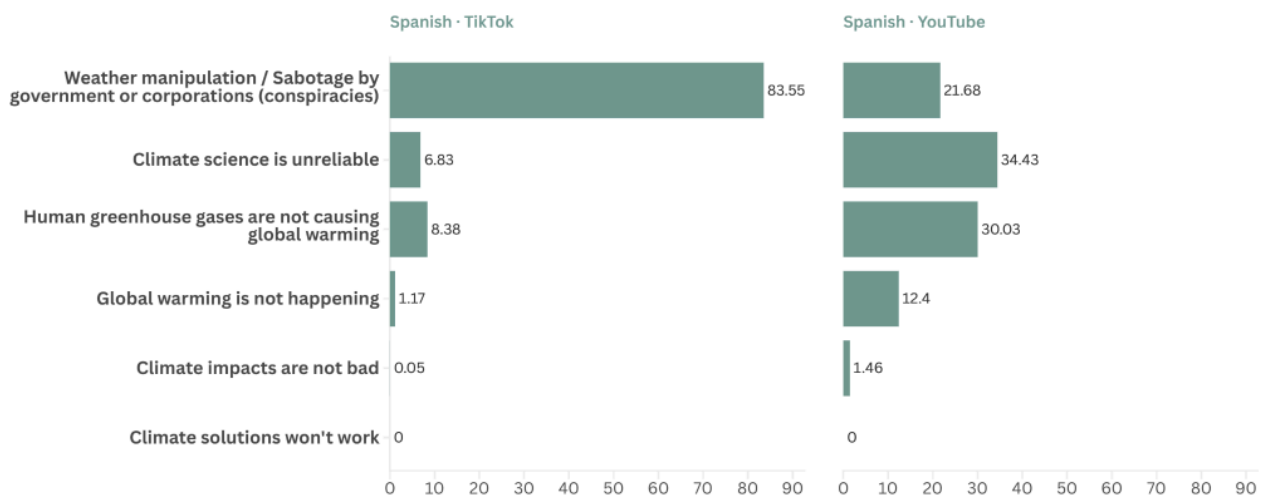
In the initial data collection of climate dis/misinformation videos, **weather-manipulation and chemtrails conspiracy narratives were most prevalent on TikTok**, with over 995,000 views in Spanish and over 600,000 in French, far outpacing more traditional denial narratives such as "climate science is unreliable", "human greenhouse gases are not causing global warming", "global warming is not happening", or "climate solutions won't work". Recent research has termed this shift the "new denial", an expansion of climate disinformation from narratives denying the existence of global warming to more sophisticated tactics, including arguing that climate solutions won't work, that the effects of global warming are in fact positive or harmless, or spreading conspiracy theories that question the motivations behind the climate movement and its policies. These "new denial" narratives do not always fit neatly into platforms' climate dis/misinformation policies, which tend to miss subtler forms, for instance, the misleading claim that CO₂ is beneficial for plants or conspiracy theories and pseudoscience that offer alternative explanations for global warming.

On YouTube, in Spanish, despite weather manipulation being the most recurrent by volume (number of videos), narratives portraying climate science as unreliable gathered most of the views of the sample. In French, the most prevalent narrative, climate science portrayed as untrustworthy or conspiratorial, was nearly half the dataset, but weather-manipulation content, despite being less common, still accumulated the most views overall (see Figure below).

Share of Views by Narrative

% of total views within each language-platform sample





Phase 1: Enforcement of platform policies

YouTube does not enforce its climate policies

Lack of transparency around enforcement actions:

The analysis across both languages shows that YouTube's ad policy enforcement is practically non-existent. Only 4-5% of the videos that did contain climate misinformation in violation of the platform's policy that we reported were seemingly demonetised following our reports. Moreover, neither Fundación Maldita.es nor Science Feedback received any response to their over 80 reports, denoting that this amounts to a structural problem in terms of transparency around enforcement action across different markets rather than a country-specific one. YouTube itself, in [its reporting page](#), only lists removal or age restrictions as actions that can be taken against reported content.

Some of these videos presented information panels below, linking to authoritative sources like the United Nations. According to its [2025 DSA risk assessment](#), these panels appear on "videos related to topics prone to misleading information". However, YouTube continues to monetise videos that contain climate disinformation. Our analysis confirms that, as do [other studies](#) that [echo the findings](#).

TikTok's enforcement is inconsistent and likely fully automated

Inconsistent enforcement:

TikTok did restrict roughly a fifth of the reported videos in each language across the initial collection and repeat offender round, but enforcement was inconsistent. Some content relating to chemtrails and geoengineering was restricted while a large share of similar claims in both languages were left unmoderated.

Even if TikTok is one of the few platforms with a dedicated policy against climate dis/misinformation, the platform also failed to take any action against videos rejecting the reality of climate change, misrepresenting its causes, or contradicting its established environmental impact. If moderated under this policy, these claims would have warranted removal, a stronger enforcement response than the climate conspiracy claims that TikTok merely demoted.

A 0% appeal reversal rate:

None of TikTok's original decisions in either language were overturned by appeals. This 0% reversal rate on 147 pieces of content raises serious questions about whether TikTok's appeals procedure is either a legal formality or a real mechanism to reverse decisions where the complainant shows they are not justified by platforms' policies.

Response time and automation:

The average response time on TikTok in France was more than double that of Spain. Spain's response pattern showed clear signs of automation, landing precisely on the half-hour or hour every time. France's responses mostly followed the same pattern but included enough exceptions to suggest at least some human review. This is consistent with TikTok's [latest transparency report](#), where they note that 'automated systems actioned 94.1% of violating content without human review.' Nevertheless, TikTok notice and action interface showed recurrent bugs on the date/time of reports, responses and appeals (see Annex I) which hindered the evaluation of the system.

Phase 2: Account-level enforcement

Repeat offenders face no consequences on both platforms

Account-level enforcement is nonexistent across all languages and platforms

Accounts that gathered at least three violative videos under platforms' policies showed no visible enforcement signals, despite the fact that both platforms say they enforce a strike policy. Evidence in this phase was limited on two fronts. First, few accounts initially reached the 3-strike level due to platforms' own lax moderation of climate dis/misinformation. Second, on YouTube in particular, it was difficult to conclusively determine if an account's violative content had been demonetized, given that YouTube did not respond to our reports and the broader lack of transparency around enforcement actions (see Limitations section for more detail).

These findings are unsurprising given that in TikTok's latest [transparency report](#), service suspension actions under their Misinformation policy are 404 out of more than 1.2M actions. That represents 0,03 %. In the case of [Youtube](#), they claim in the report to have taken 153,000 actions under the advertiser-friendly content guidelines, which includes climate dis/misinformation. In all cases, it led to monetary restrictions, but again, due to the lack of transparency around monetisation data, it was impossible to discern with certainty whether any monetary restriction had been applied to accounts with various reports of violative content.

Phase 3: Appeal mechanism

Appeal mechanism

Out-of-Court-Dispute-Settlement (ODS)

One month after filing our last complaint with the ODS, we had received resolutions on just 10% of our complaints. The ODS response rate for TikTok was 93% and for YouTube 7%, with an additional 56% of complaints resolved across both platforms at the time of publication. The overturn rate, meaning the rate at which the appeals body recommended reversing the platform's original decision, was 61%. In an additional 16%, the ODS decided in our favour on procedural grounds, because the platform had failed to provide enough information to make a decision, and in the remaining 23%, the ODS sided with the platform's original decision.

Digital Services Coordinator (DSC)

In France, we filed two complaints with Arcom, the French DSC, under Article 53 DSA, one concerning TikTok and one concerning YouTube. Arcom cannot assess the lawfulness of content or require a platform to reverse a moderation decision. The complaints therefore asked it to examine whether the platforms apply their own rules on climate dis/misinformation diligently (Art. 14) and handle reports and appeals properly (Art. 20). Each was accompanied by a database of the reported videos and the platforms' responses. Article 53 provides that complainants receive appropriate information on the status of their complaint. At the time of writing, we had received no reply. We will update this report if and when we hear back.

3. Findings by Language and Platform

3.1 FRENCH

- TIKTOK

PHASE 1:
Enforcement of
platform policies

- From the videos Science Feedback found to be violative of TikTok's policies, the platform only took action and restricted 3% of them, all under their rule against conspiracy theories. Moderation was not consistent since TikTok left 27 other videos that also contained similar conspiracy claims relating to chemtrails/geoengineering unmoderated.
- Conspiracy theories related to narratives about weather manipulation and geoengineering (e.g., chemtrails, HAARP) made up 54.7% of all videos and accumulated over 600,000 views. Moreover, 80% of the ten most-viewed videos included conspiracies on weather control or conspiracies that imply climate change is used by governments to instill panic and control. For instance, [this video](#) suggesting governments produce climate change through [secret technologies](#) accrued over 500.000 views. Some of the top ten most viewed videos on TikTok showed signs they have been amplified by the platform's recommendation algorithm, with one video reaching more than 10 times its account's follower base and another reaching over 8 times.
- TikTok upheld its decision in each case during the appeal round; in other words, the appeal reversal rate was 0%.
- Response times averaged around 38 hours, with 30% of responses landing on an exact hour or half-hour mark, indicating an automated response. For the remaining 70%, the first-response date shown in TikTok's reports page is identical to the date of our appeal and, in some cases, falls after the final appeal decision. This appears to be a bug in TikTok's interface, which overwrites the first-response date with the appeal date, so TikTok's actual response time cannot be determined for these cases. See Annex I for details.

PHASE 2:
Account level
enforcement

- We tested reporting various videos from repeat offender accounts to try to test account-level enforcement, but even for accounts with at least 3 strikes and over 134 followers, there were no visible account-level actions

PHASE 3:
Escalation
through DSC

- Under Article 53, we filed a complaint with Arcom concerning TikTok, covering 101 videos for which TikTok upheld its decision on appeal. At the time of writing, we had received no reply. We will update this report if and when we hear back.

- YOUTUBE

PHASE 1:
Enforcement of
platform policies

- YouTube provided no response to the reports of violative content submitted to the platform under its policies. Therefore, we relied on tools that gather monetisation signals and on manually reviewing ads before playback (see methodology for more details) and found that only 5% of the reported videos showed signs of having been demonetised 72 hours after reporting, suggesting possible action taken. 2% was deleted, while the remaining 93% still showed signs of being monetised.
- Climate science being portrayed as untrustworthy or conspiratorial made up 47,1% of all videos and accumulated over 185,000 views. For example, [this video](#) promotes the idea that governments are manipulating public opinion and that information about [climate change is unfounded](#) and intended solely to spread panic. Weather manipulation was still the narrative cluster with the most views overall, totalling more than 300.000. For instance, [this video](#), with more than 200,000 views, amplifies false and harmful claims about large-scale, ongoing [“chemtrail” and geoengineering programs](#).
- YouTube showed few signs of algorithmic amplification, with most of the top ten videos with the most views performing poorly relative to their follower base.

PHASE 2:
Account-level
enforcement

- We reported 3 videos from a single account with 134 followers. After our reports, all three appeared to have been demonetised. The account itself, however, displayed no sign of any account-level measure, and the creator continued to post several times a day.

PHASE 3:
Escalation
through DSC

- Under Article 53 we filed a complaint with Arcom concerning YouTube. YouTube did not respond to any of our reports, so there was no decision to appeal. The complaint concerns the lack of response and the enforcement of its monetisation rules. At the time of writing, we had received no reply. We will update this report if and when we hear back.

3.1 SPANISH

- TIKTOK

PHASE 1: Enforcement of platform policies

- From the videos Fundación Maldita found to be violative under TikTok's policies, the platform only took action and restricted 17% of them, all under their rule against conspiracy theories. Moderation was not consistent since TikTok had a number of additional videos that also contained similar claims relating to chemtrails/geoengineering unmoderated.
- More than half of the videos (60%) related to conspiracy theories on weather manipulation (chemtrails, HAARP and geoengineering) or narratives of sabotage by states or corporations, accumulating over 995,000 views. Some of these videos showed signs of algorithmic amplification. For instance, [this video](#), with over 119,000 views, spreads a [weather manipulation conspiracy](#) through an unfounded claim that a weather event was deliberately manipulated for economic interests during a marathon in Valencia. This video received over double the views of its account's follower base. Similarly, [this other video](#), with over 38,000 views, spreads the [conspiracy theory](#) that wildfires in Chile were provoked by a corporation seeking to obtain mining permits for rare mineral extraction on the affected land. In this case, the view rate also more than doubled the account's follower base.
- TikTok upheld its decision in each case during the appeal round; in other words, the appeal reversal rate was 0%.
- The response time for reports averaged around 22 hours, and every response landed on an exact half-hour or hour mark, suggesting automated processing rather than individual manual review.

PHASE 2: Account-level enforcement

- To test account-level enforcement, we reported several videos from repeat-offender accounts. Even in the case of an account with more than 2,600 followers and three reported violations, we observed no visible account-level action.

PHASE 3: Escalation through ODS

- One month after we submitted the disputes to the ODS for TikTok, the appeals board has rendered a decision on 28% of the cases.
- Of these, the appeals body decided in 50% of cases that TikTok needed to apply "relevant content restrictions (for example, warning screens or other actions)", likely referring to demotion.
- In 8% of the cases, it recommended outright removal.
- In 26% of cases, the ODS found the content did "not violate the platform's policies and should remain on the platform".
- The remaining 16% resulted in a procedural ruling in our favour, because TikTok had not "provided sufficient information to allow us to make a decision regarding your case".
- In all cases, there were no visible signs the ODS' recommendations had been applied by the platform.

- **YOUTUBE**

**PHASE 1:
Enforcement of
platform policies**

- YouTube provided no response to the reports of violative content submitted to the platform under their policies. Therefore, we relied on tools that gather monetisation signals and on manually reviewing ads before playback (see methodology for more details) and found that only 4% of the reported videos showed signs of having been demonetised 72 hours after reporting, suggesting possible action taken. Another 2% could not be verified, while the remaining 94% still showed signs of being monetised.
- From these, 31% of videos contained claims on conspiracy theory narratives relating to weather manipulation or narratives linked to state-driven conspiracies of sabotage and accumulated over 133,000 views. Take, for instance, [this video](#), with over 29,000 views, that promotes the [conspiracy theory](#) that chemtrails are used to manipulate weather conditions and produce droughts and floods. Yet, despite being the third in volume, the narrative cluster “climate movement/science is unreliable” received the most views, accumulating over 200,000. For example, [this video](#), with over 20,000 views, portrays climate change as a [scam](#) designed to control people.
- YouTube showed few signs of amplification, with most of the top ten videos with the most views performing poorly relative to their follower base.

**PHASE 2:
Account level
enforcement**

- We tested reporting various videos from repeat offender accounts to try to test account-level enforcement. After reporting at least three violative videos per channel, according to platform policies, we monitored whether creators continued posting over the following two weeks as a proxy for whether any account-level restriction had been imposed.
- Only 1 account showed a pattern consistent with a possible restriction, though this cannot be confirmed with certainty given the platform's lack of disclosure on moderation decisions (see section on limitations and recommendations).

**PHASE 3:
Escalation
through ODS**

- One month after we submitted the disputes to the ODS for YouTube, the appeals board has rendered a decision on 1% of the cases.
- For these, the ODS body agreed in all cases that the content should stay up and also noted the content did “meet the criteria for being restricted under the applicable policy, and the platform must continue to apply the relevant content restrictions (such as information banners or other actions)”.
- In all cases, the video presented an information banner but still showed signs of being monetised.

4. Recommendations

The project set out to test compliance with Digital Services Act (DSA) Articles 14, 20, 21, 34 and 35 by examining the application of terms and conditions, complaint handling, out-of-court dispute settlement, and risk mitigation. It also assesses self-regulatory adherence to the EU Code of Conduct on Disinformation, of which YouTube and TikTok are signatories, including the commitments to reduce the amplification of misleading content and demonetise disinformation. As such, findings can inform regulators on monetisation of disinformation, the effectiveness of enforcement mechanisms, and platform accountability overall.

Recommendations for Platforms

→ **Apply existing policies diligently and consistently (Art. 14 DSA).**

If a platform sets out content rules against climate dis/misinformation, or its monetisation, but fails to apply them consistently, it is falling short of its own obligations. Intermediary services must apply and enforce their policies diligently, impartially, and proportionately in accordance with Article 14 of the DSA. As such, both platforms' enforcement of their policies regarding climate dis/misinformation is mandatory and serves as one of the benchmarks to measure DSA compliance.

→ **Update climate policies to reflect "new denial" narratives.**

Climate policies are too narrowly defined. In the case of YouTube, they only encapsulate traditional denial narratives, for instance, denial of scientific consensus, denial of human-induced greenhouse effects, or outright denial that global warming is happening. They fail to account for 'new denial' narratives such as those framing climate change as beneficial for nature; conspiracies that portray the scientific community as biased or fraudulent; claims that governments manipulate weather conditions through geoengineering; or that global warming is a hoax, made up by elites to advance a globalist agenda. Platforms should update their policies to evolve alongside dis/misinformation narratives. For platforms like TikTok that do account for these newer narratives, enforcement should be consistent and not limited to a fraction of content that carries lesser penalties like demotion.

→ **Ensure human review meaningfully improves automated moderation (Art. 20 DSA).**

Under Article 20, platforms must create an internal appeals process with the ability to hear appeals across content types, violation types, and languages. They must do so under "the supervision of appropriately qualified staff, and not solely on the basis of automated means".

While it would be unreasonable to expect a moderator to review every piece of content, platforms should ensure human review actually improves the performance of automated moderation systems overall, precisely to prevent inconsistencies such as near-identical content being moderated differently, or a 0% appeal-reversal rate. These point to a systemic reliability gap in TikTok's moderation, not isolated cases.

Recommendations for Regulators

→ **Ensure platforms comply with Articles 34 and 35 of the DSA.**

Articles 34 and 35 place an obligation on Very Large Online Platforms to identify, assess, and mitigate systemic risks arising from their services. Disinformation is explicitly listed among those risks. Yet climate dis/misinformation intersects with other systemic risks like public health and safety. For instance, videos claiming that CO₂ is harmless, or [false narratives](#) that spread in the [aftermath of natural disasters](#). Platforms are also required to adapt and rectify failures when mitigation measures prove insufficient. Our evidence suggests YouTube's ad revenue architecture and TikTok's algorithmic amplification facilitate climate-related systemic risks rather than mitigate them.

→ **Ensure platforms comply with their transparency obligations (Art. 20 DSA).**

The absence of a reply from YouTube to our in-platform reports of violative content under their policies goes against Article 20 of the DSA. This article obliges platforms to inform complainants of whether or not a moderation action was taken upon the receipt of their notice, along with the available options for redress. This mechanism is meant to give users and researchers visibility into how platforms apply, or in the particular case studied, fail to apply, their own rules in practice.

→ **Ensure signatories comply with the Code of Conduct on Disinformation.**

- Under the Code of Conduct, signatories should ensure they are taking appropriate measures to prevent the financing of purveyors of disinformation. YouTube's Partner Program [pays creators 55%](#) of net revenues from ads displayed on their videos, with YouTube retaining the remaining 45%. This means that every view on a climate misinformation video generates income for both the creator and the platform itself. As such, YouTube not only hosts and effectively finances this content; it also financially benefits from it. This goes hand in hand with requiring greater transparency around monetisation. At present, Meta is the only platform that publicly provides some meaningful information about the monetisation of content and accounts. Transparency on monetisation should therefore be available across all platforms and include data on what accounts participate in revenue-sharing programs, how revenue is shared across creators, create specific pathways for demonetisation in internal complaint-handling mechanisms, and inform users when a report results in demonetisation of the violative content.
- TikTok is a signatory of the fact-checking chapter of the Code and should, accordingly, ensure consistent fact-checking coverage on its platform. The company emphasized its collaboration with independent fact-checking partners and climate experts in its latest [risk assessment](#). However, our findings indicate otherwise. Not only did we find videos containing climate dis/misinformation available on the platform, but none of them had labels. Arguably, this implies that TikTok can make better use of its cooperation with fact-checkers and scientific experts, at least when it comes to climate content, and indicates a more general lack of commitment to its obligations under the Code.

Recommendations on Out-of-Court Dispute Settlement Bodies (Art. 21 DSA)

Article 21 of the DSA states that users are entitled to select any out-of-court dispute settlement body (ODS) "to resolve [...] complaints that have not been resolved by means of the internal complaint-handling system".

→ **Limitations to standardised response templates.**

ODS bodies rely on response templates that do not always reflect the reality of the case. For instance, in one of our reports, the ODS found the content should not have been removed, yet when we reported the video, TikTok had found no violation to begin with, and the video remains visibly online today at the time of writing. The same lack of granularity pitfall appears when submitting a new dispute. The form provides no option beyond "the platform left up content that I want removed", and does not allow reporting parties to request, for example, that content be restricted or demoted rather than removed. This binary framing does not fit all content types, especially something as nuanced as dis/misinformation.

→ **Improve response times.**

Only 10% of our reports have obtained resolutions in the month since we submitted them to the ODS body. The Commission and Digital Services Coordinators (DSCs) should remind platforms that, under Article 21 of the regulation, they must act promptly and in good faith to provide ODS bodies with enough information so they can respond to reports promptly, but ODSs themselves must account for this excessive period before users get a reply.

→ **Improve visibility of enforcement below the threshold of removal.**

One key limitation of the system is that recommendations issued below the threshold of removal have no visibility on platforms. There is no way to monitor or track whether ODS recommendations are actually taken up by platforms. Given these difficulties, regulators must continuously assess whether the platforms' engagement with the bodies is consistent with the principle of good faith established in the regulation. Although ODS decisions are not binding, platforms should be expected to explain why they do not implement them, particularly where they fail to do so at scale. As [Husovec aptly puts it](#), "the effect of ODS decisions is best described as 'systemically but not individually binding' ". This implies that while platforms can dismiss specific requests, they shouldn't be able to consistently overlook the mechanism as a whole without justification.

5. Methodology

DATA COLLECTION

To collect data, we relied on a bilingual (French/Spanish) list of keywords that are frequently found in climate disinformation posts based on several years of monitoring by Fundación Maldita.es and Science Feedback (see the keyword list Annex II). Using third-party API access to YouTube and TikTok, we collected videos containing these keywords published from October 2025 to February 2026. The results of this data collection were 14,000 videos on TikTok (Spain: 4,500, France: 9,400) and 70,000 videos on YouTube (Spain: 29,000, France: 41,000).

On the description and hashtags of each video, we applied a large language model (LLM, Mistral) as a topic-relevance filter, retaining only videos considered relevant to the scope of the study and removing non-relevant content, such as entertainment. This filtering resulted in 4,400 relevant videos on TikTok and 20,000 videos on YouTube.

For videos for which a transcript was available, we subsequently used an OpenAI model to assess their content against the relevant platform policies, using a prompt previously tested by Science Feedback. Each video was assigned a score ranging from 0 (no indication of a policy violation) to 10 (clear indication of a policy violation); only videos receiving a score above 7 were further considered. This pre-screening enabled scalable identification of likely violative content under platform policies and resulted in 331 TikTok videos (Spain: 133, France: 198) and 894 flagged YouTube videos (Spain: 564, France: 330) being flagged as likely violative. A further step was to remove all of the posts that had under 100 views from analysis. This ensured that we were reviewing and reporting content that was reaching a wider number of users. Fact-checkers in both organisations then manually reviewed the highest-impact videos against debunked claims by Fundación Maldita.es, Science Feedback, or another authoritative source. After manually reviewing each piece of content and fact-checking it, from this initial data collection, Fundación Maldita.es identified 35 violative videos according to TikTok's policies and 48 under YouTube's policy, and Science Feedback identified 52 violative videos under TikTok's policies and 34 videos under YouTube's policy. Each confirmed case was recorded in a database with a short explanation of why the content constituted climate dis/misinformation and was violative of platforms' terms; fact-checkers also added a link to a fact-check or authoritative source rebutting the misleading claim(s).

NARRATIVE CLUSTER ANALYSIS

Narrative clusters were identified using the taxonomy of [climate disinformation narratives](#) developed by Coan et al., which organises claims into super-claims and sub-claims. We added a sixth super-claim, "Conspiracy: weather manipulation/sabotage by government or corporations", to capture content that falls under the "new denial" narratives identified earlier but is not included in the original taxonomy.

To assess whether certain narrative clusters were more amplified than others, we took the 10 most-viewed videos on each platform and calculated their views relative to the creator's follower base to understand whether the reach of this content was likely organic or likely driven by algorithmic amplification.

PHASE 1: ENFORCEMENT OF PLATFORM POLICIES

Violative videos under platform policies were then reported through each platform's internal complaint-handling systems.

- **TIKTOK**

Videos were reported under the 'Harmful Misinformation' category, per its [Integrity and Authenticity policy](#), which warrants the takedown of any "misinformation that denies the existence of climate change, misrepresents its causes, or contradicts its established environmental impact". The same policy should also cover adjacent content types that often overlap with climate disinformation, namely conspiracy theories, unverified claims during a crisis or major event, or tactics such as cherry-picking of scientific data. However, there is a separate, more general policy warranting a lighter enforcement action for them, that of demotion rather than takedown. This enforcement gap let us infer which policy TikTok actually applied to our reports. Whereas a takedown would signal enforcement under the climate policy itself, demotion signaled enforcement under an adjacent, less restrictive one.

- **YOUTUBE**

We only selected monetised videos, since demonetisation is the only enforcement pathway the platform sets out for climate dis/misinformation (see section on limitations). Its [advertiser-friendly guidelines](#), under 'Harmful acts and unreliable content,' state that content "contradicting scientific consensus on the existence of and causes behind climate change" will receive no ad earnings. As per this policy, we reported the videos under the 'misinformation' category.

PHASE 2: ACCOUNT-LEVEL ENFORCEMENT

For accounts that had received a moderation decision, we manually searched for other violative videos under platforms' policies and submitted additional reports over the course of the project period to observe for account-level enforcement. We prioritised "worst offenders", meaning accounts already known to both organisations as recurrent disinformers with a history of repeated policy violations. Searching these accounts specifically, we identified an additional 26 violative videos on TikTok and 36 on YouTube in Spanish, and 61 additional videos on TikTok and 45 on YouTube in French. We then tracked platform responses to these reports to determine whether they resulted in content-level or account-level action, monitoring specifically for accounts with 3 or more violative videos.

- **TIKTOK**

Its [strike policy](#) does not specify a fixed threshold for account bans, as its strike system varies by the severity of the violation. We considered it reasonable, however, to treat 3 strikes as the threshold warranting account-level enforcement.

- **YOUTUBE**

Its [strike policy](#) is explicit: creators who receive 2 strikes within 90 days of a first strike "won't be allowed to post content for 2 weeks". Given this clear threshold, for creators with 3 or more reported videos, we monitored their account for 2 weeks following the report to check for any visible restriction, taking into account the channel's usual upload frequency.

This was the most difficult enforcement pathway to measure, given the platforms' lack of transparency around whether and when individual reports translate into account-level action (see limitations section).

PHASE 3: APPEAL MECHANISM

In cases where platforms found our reports not to be violative of their policies, even after an internal appeals round, we escalated these cases, drawn from both the initial data collection and the "worst offenders" round, through the DSA's contemplated external appeal mechanisms. In France, this meant filing with ARCOM, the Digital Services Coordinator. In Spain, with [Appeals Centre Europe](#), an Out-of-Court Dispute Settlement (ODS) body certified to review cases relating to the application and enforcement of terms and conditions of social media online platforms. At the time of writing, roughly one month had passed since the Spanish content was appealed, so the figures reported here reflect only the responses received from the ODS within that period. In France, we filed one complaint per platform with Arcom, the Digital Services Coordinator, under Article 53 DSA, which lets recipients of a service lodge complaints alleging an infringement of the Regulation. As Arcom cannot rule on content or overturn moderation decisions, the complaints were framed around the platforms' conduct: application of their own terms (Art. 14), and handling of reports and appeals (Art. 20). At the time of writing, we had received no reply. We will update this report if and when we hear back.

LIMITATIONS

- **Monetisation on YouTube**

To determine the monetisation status of YouTube videos, we used an automated detection system to identify the presence of ads on confirmed violative videos under YouTube's policies. We then confirmed this manually using two different [third-party monetisation checkers](#), cross-checking with the [TubeLab](#) and [Nexlev](#) extensions and documenting whether ads appeared before playback to reasonably infer whether a video was monetised. We then compared the monetisation status before and after reporting, which allows us to understand whether an action could have been taken as a result of our report. However, these results are not fully conclusive, given YouTube's lack of transparency around monetisation data. Even if YouTube's penalty for climate misinformation is demonetisation rather than removal, its [reporting page](#) only lists removals and age restrictions as possible outcomes of a report. Demonetisation does not appear as a moderation action at all. As a result, when a report leads to the only penalty the policy actually foresees, the complainant has no way of knowing whether it was enforced or not.

- **Account-level enforcement**

Account-level enforcement was by far the most difficult outcome to test. First, platform policies are unclear about their strike systems and about what number of violative videos under the policy actually triggers account-level action. Second, many forms of account-level restriction, such as demonetisation or demotion, aren't visible from the outside, making it impossible to accurately track anything short of an outright account ban. This was even more difficult on YouTube, where the complete lack of transparency around enforcement actions tied to demonetisation meant we had no reliable way to track strikes at all.

Please refer to the recommendations section for appropriate actions to address these shortfalls.

ANNEX I

TikTok User Interface Bugs

Data from TikTok's reports page was entered manually into an external database, recording for each video the date of the report, TikTok's first response, the appeal, and the response to the appeal. Some entries were made step by step as the process unfolded; others were made once it was complete. This is how we discovered that, once an appeal is filed, TikTok's reports page replaces the date of its first response with the date of the appeal. In Figure 1, the first response and the appeal carry the same timestamp, down to the second, which cannot be correct. In some cases, the page even shows the date of the appeal before a first decision was made by TikTok, which chronologically is not possible (Figure 2).

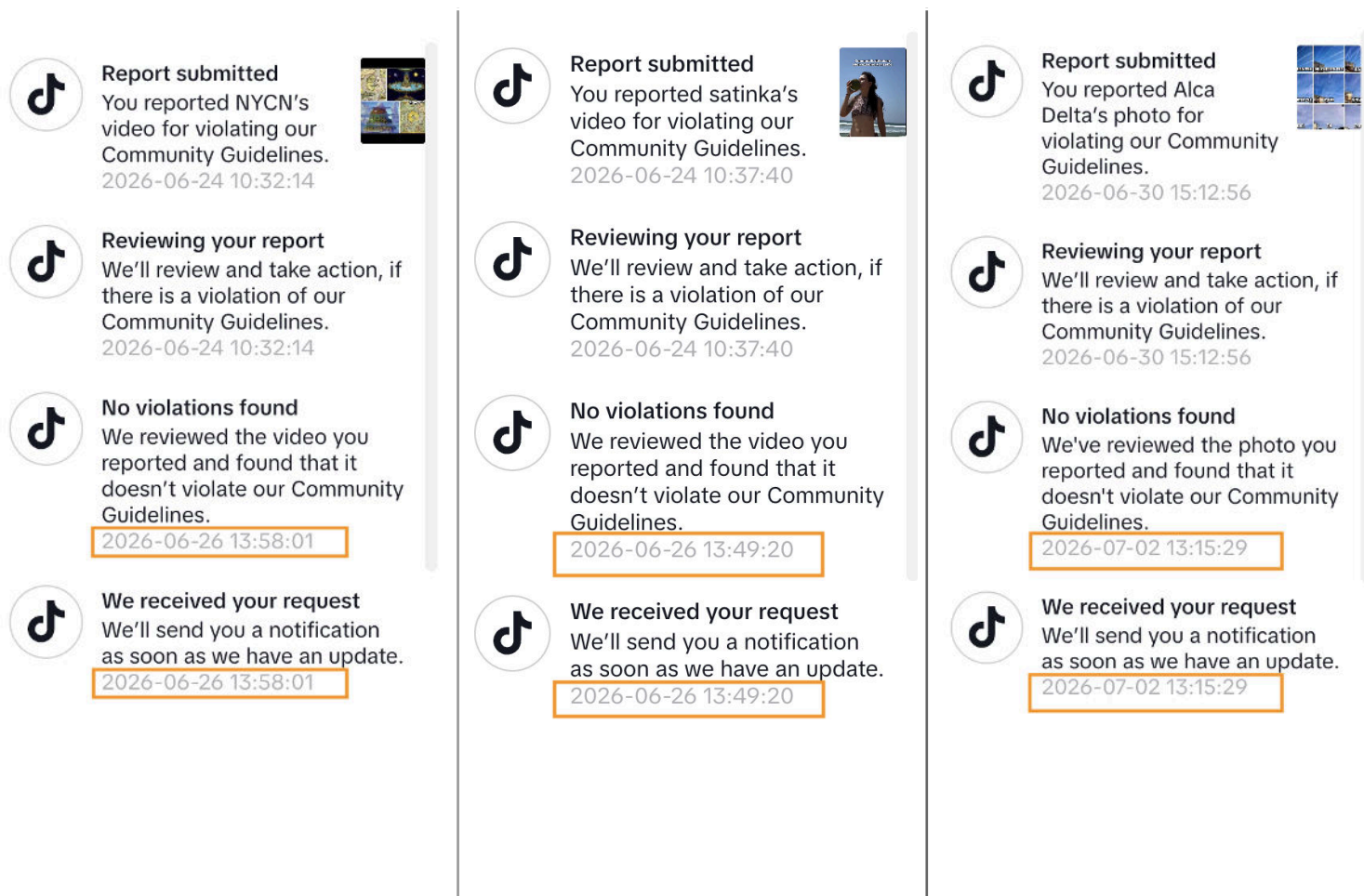


Figure 1. Screenshots of TikTok interface where date and time of TikTok's first response was overwritten with the appeal date and time.

TikToks URL:

- (1) <https://www.tiktok.com/@nycn67/video/7647466232270032160> ,
- (2) <https://www.tiktok.com/@satinkashowdanse/video/7613906566898552086> ,
- (3) <https://www.tiktok.com/@alca.delta/photo/7653860955926875425>

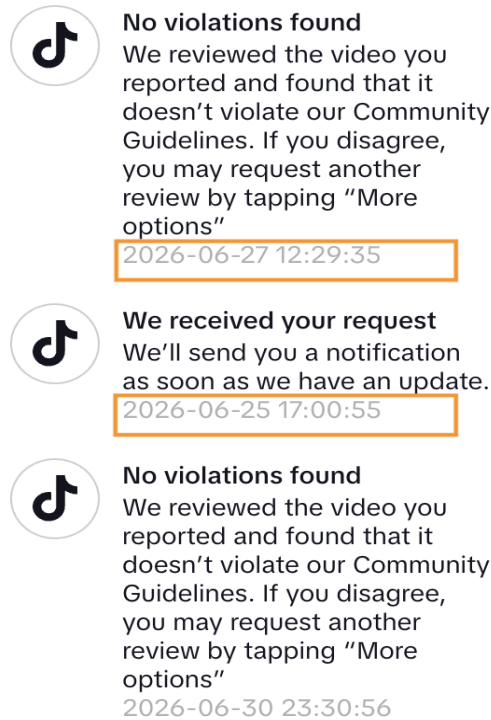


Figure 2. Screenshot of TikTok interface where the appeal date appears before the first response of the platform.

TikTok URL: <https://www.tiktok.com/@alca.delta/video/7650741962491858208>

We therefore treat the first-response date as unreliable for any video whose entry was completed after the appeal, and base the response-time analysis only on the 18 videos where it was recorded before the appeal or where no appeal was filed. For these, TikTok responded after an average of about 6h. For the remaining 34 videos, the average response is 55h, although we cannot be sure if these videos were actually manually assessed or are just a byproduct of TikTok's bug.

ANNEX II

Keywords List

Keywords in Spanish			
(China OR India) AND cambio climático	ciclos naturales AND cambio climático	fraude climático	negocio AND placas solares
(China OR India) AND emisiones	ciclos solares	fumigaciones masivas	No es sequía, es saqueo
97 por ciento AND CO2	ciudad 15	geoingeniería	no existe AND cambio climático
97 por ciento AND dióxido de carbono	ciudad de 15 minutos	Geoingeniería Marruecos AND	No tendrás nada y serás feliz
Agenda 2030	ciudades inteligentes AND cambio climático	geoingeniería militar	Nubes estáticas
Agenda climática	clima cambiático	globalista	Nuevo Orden Mundial
Agenda Oculta	CO2 y volcán	globalistas	nuevo orden mundial AND cambio climático
alarmismo AND cambio climático	CO2 y volcanes	globalistas del clima	olivos AND placas solares
alarmista AND cambio climático	coche eléctrico AND cambio climático	gobierno mundial AND agenda 2030	Otanistas
alarmista climático	coche eléctrico AND CO2	Greta Thunberg	pájaros generadores eólicos
apocalipsis climático	coche eléctrico AND estafa	guerra electromagnética	pájaros molinos eólicos
armas climatológicas	coche eléctrico AND incendio	gurús del clima	presas AND cambio climático
armas geofísicas	confinamiento climático	HAARP AND cambio climático	prohibición de combustibles
ataque climático	control AND cambio climático	histeria climática	radares meteorológicos
ataque climatológico	destrucción de presas AND cambio climático	Hunga-Tonga	religión AND cambio climático
ataques climatológicos	dióxido de carbono AND natural	iglesia de la calentología	religión climática
ataques meteorológicos	dogma AND cambio climático	impulsos electromagnéticos AND cambio climático	religión progre AND cambio climático
aviones dibujando líneas	ecofascistas	impulsos electromagnéticos AND cambio climático	sales de bario AND lluvias
ballenas generadores eólicos	ecoidiota	incendios AND molinos eólicos	sequía inducida
ballenas molinos eólicos	ecolojetas	incendios AND renovables	sequía provocada
calentamiento artificial	ecolojetas	incendios AND tierras raras	Siembra de nubes
calentamiento asintomático global	efecto invernadero inducido	ingeniería del clima	terroristas climáticos

calentología	élites globalistas AND cambio climático	intervención climática	Timo cambiático
calentón global	embalses AND sequía	jet privados AND cumbre del clima	Timo climático
calimas	engaño climático	Ley de Aguas	timo cromático
cambio cromático	estafa AND cambio climático	limpieza de montes AND incendios	tiranía AND agenda 2030
camelo climático	estafa climática	lobby climático	vamos a comer placas solares
campaña antigranizo	estelas	mafia 2030	Verdad AND cambio climático
chemtrails	estelas AND aviones	manipulación climática	verdad oculta AND cambio climático
China AND India AND cambio climático	estelas químicas	manipulación del clima	woke AND cambio climático
China AND India AND emisiones	falacia del clima	mentira climática	yoduro de plata
chiringuito AND cambio climático	farsemia climática	modificación climática	zonas de bajas emisiones AND cambio climático
ciclos de milankovitch	fraude AND calentología	negocio AND energías renovables	zurdos AND cambio climático

Keywords in French			
Agenda climatique	mensonge climatique	HAARP AND changement climatique	gouvernement mondial AND "agenda 2030"
Alarmiste climatique	terroristes climatiques	manipulation du climat	sels de baryum
Alarmisme climatique	Géoingénierie	manipulation climatique	guerre électromagnétique
Agenda 2030	Interdiction AND moteurs à combustion	élites mondialistes	traînées AND avions
Apocalypse climatique	refroidissement climatique	armes climatologiques	traînées chimiques
arnaque climatique	canular climatique	hystérie climatique	dioxyde de carbone AND "naturel"
Nouvel Ordre mondial	imposture climatique	fraude climatique	religion climatique
ensemencement des nuages	farce climatique	alarmiste AND changement climatique	religion AND changement climatique
chemtrails	#ClimateScam	alarmisme AND changement climatique	modification climatique
escroquerie climatique	#ClimateHoax	cycles naturels AND	CO2 AND "volcan"

baleines éoliennes	#Geoengineering	GIEC arnaque	CO2 AND "volcans"
écofascistes	#Chemtrails	climatologues corrompus	géo-ingénierie
réchauffement artificiel	#ClimateTruth	Climato-réaliste	mondialiste
bobard climatique	modèles climatiques bidon	climatorealiste	mondialistes
ville de 15 minutes	climategate	Idriss Aberkane AND climat	globaliste
éco-idiot	pseudo-science climatique	Vincent Lapierre AND climat	globalistes
fanatique AND climat	catastrophisme climatique	canular climatique	oiseaux éoliennes
extrémistes AND climat	optimum climatique	cycles naturels AND réchauffement	changement climatique
effet de serre induit	panique climatique	réchauffement climatique AND "arnaque"	dogme climatique
tromperie climatique	exagération climatique	GIEC corruption	réchauffologie
mondialistes AND climat	débat climatique interdit	GIEC mensonge	woke AND "changement climatique"
gourous AND climat	interdiction essence	vérité AND changement climatique	pas de changement climatique
église de la réchauffologie	vérité cachée AND changement climatique	confinement climatique	armes géophysiques
lobby climatique	gauchistes AND changement climatique		