

**Tracking
Exposed**

EU Disinfo Lab conference 2022

How do we monitor algorithms?

Marc Faddoul, co-director

marc@tracking.exposed

Why we need algorithmic audits?

Recommendation algorithms are:

- the gatekeeper of the internet
- opaque and unaccountable
- not optimized for user's interests

Content Moderation

The image shows an iceberg floating in the ocean. The tip of the iceberg is above the water line, while the much larger, submerged part is below. The text 'Content Moderation' is written in white, bold, sans-serif font across the top of the image, positioned over the visible tip of the iceberg.

Algorithmic Amplification

The text 'Algorithmic Amplification' is written in white, bold, sans-serif font across the middle of the image, positioned over the submerged part of the iceberg.

Algorithmic Demotion

The text 'Algorithmic Demotion' is written in white, bold, sans-serif font across the bottom of the image, positioned over the submerged part of the iceberg.

Auditing: Cooperative VS Adversarial

Cooperative Auditing

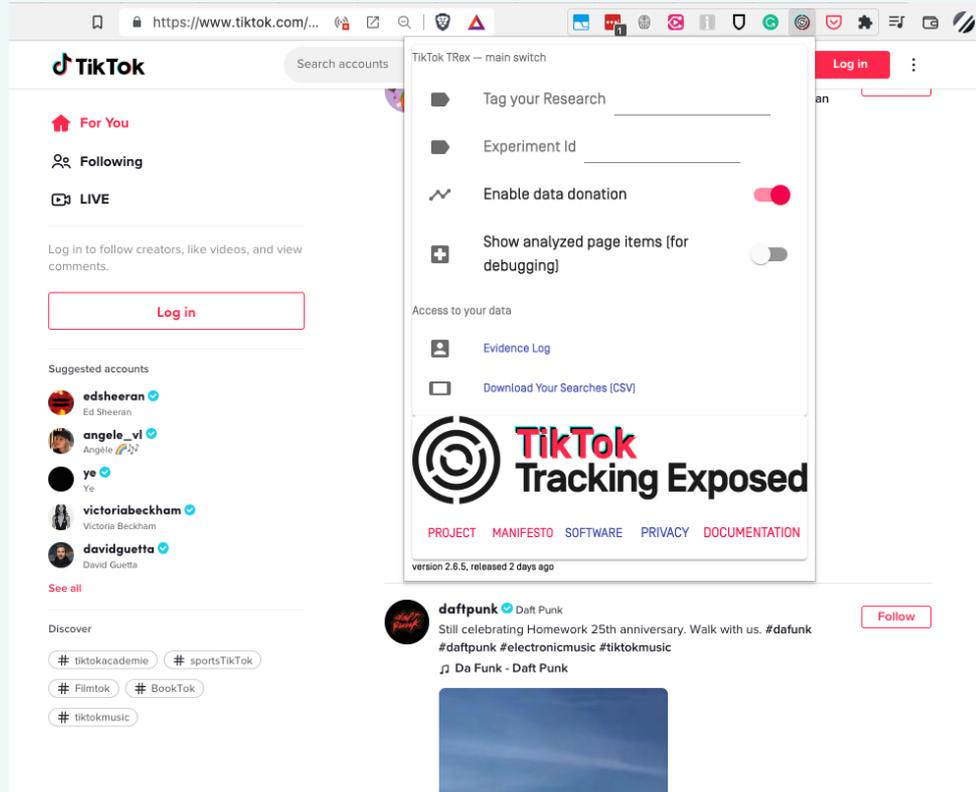
- Relies on official APIs and vetted access
- Currently very limited
- Platform initiative or imposed by DSA art 40
- Pre-defined data set and research question

Adversarial Auditing

- Data collected without platform consent
- Main driver of algorithmic accountability
- Labor and expertise intensive
- Limited to observable data

Recording: Browser extension

A browser extension to record the platform's behavior



Interacting: Puppet Accounts

A browser automation tool
to run controlled-
experiments

The screenshot displays the Guardoni browser automation interface. At the top, the logo "Guardoni" is shown with the version "-v2.4.3 - production-". Below the logo, there are social media icons for YouTube, Spotify, and a smiley face. The main interface is divided into several sections:

- Experiment:** Shows a unique ID "37384a9b7dff26184cdea226ad5666ca8cbbf456" and a link icon with the number "6". There are "CLOSE" and "RUN" buttons.
- Output:** Contains two informational boxes:
 - The first box provides a "Directive to URL" with the text: "https://www.youtube.com/watch?v=8LnZbAsws20, Loading delay 3000 [--load optional]".
 - The second box provides a "Run experiment with config" with a detailed JSON-like configuration: `{ chromePath: '/usr/bin/google-chrome', headless: true, verbose: false, loadFor: 3000, basePath: '/home/ascariandrea/tracking-exposed/trex/platforms/guardoni', profileName: 'default', evidenceTag: 'no-tag-12352', advScreenshotDir: undefined, excludeURLTag: undefined, yt: { name: 'youtube', backend: 'https://youtube-tracking-exposed/`
- YouTube Player:** A video player is embedded, showing a news segment from Fox News. The video title is "Naomi Seibt Talks Climate Realism with Dana Perino on Fox News Channel". The video has 91K views, 2.5K likes, and a "DISLIKE" button. There are also "SHARE" and "SAVE" buttons. The channel name "The Heartland Institute" is visible with 38.1K subscribers and a "SUBSCRIBE" button.



Adversarial Audit Example

TikTok's policy response in Russia
following the war in Ukraine



March 6th - First Report

Following the invasion of Ukraine, we uncovered that TikTok had blocked access to 95% of content in Russia (all international), without announcing it.

THE
WALL STREET
JOURNAL

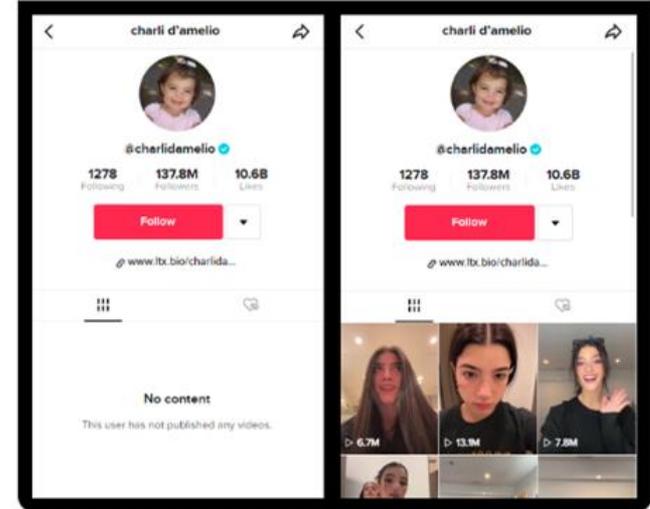
VICE News

FT FINANCIAL
TIMES

The
Guardian

TikTok as seen
from Russia

TikTok as
seen from
the rest of
the World



3:05 PM · 15 mars 2022 · Twitter Web App

404 Retweets 85 Tweets cités 721 J'aime



@trackingexposed

April 13th - Second Report

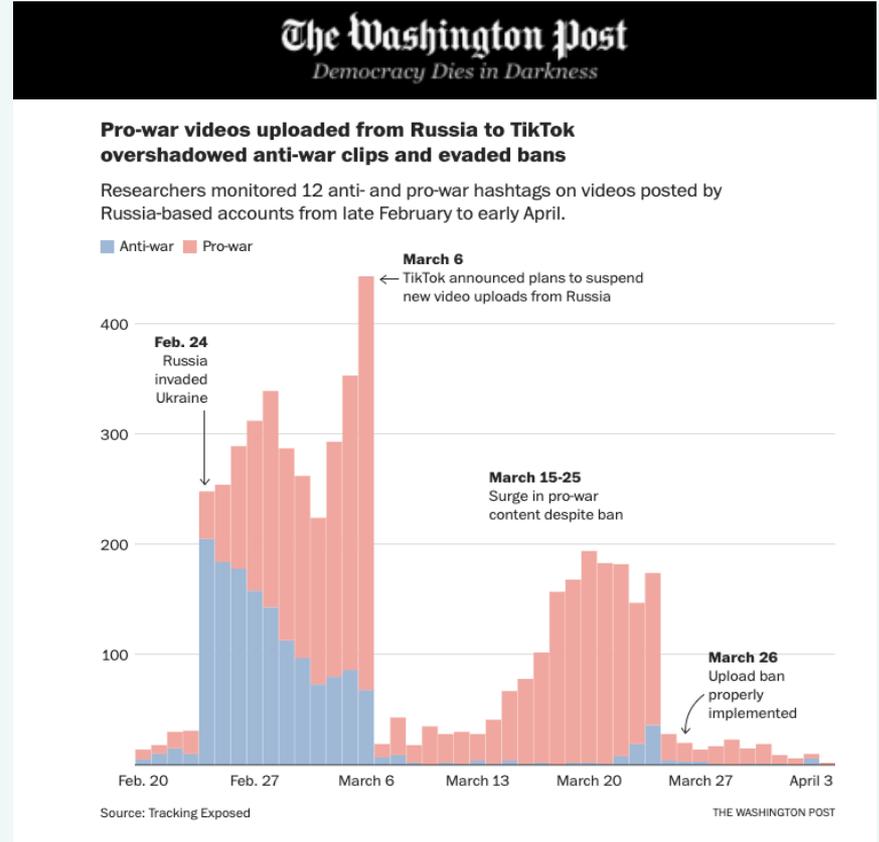
We found a loophole being exploited by pro-kremlin propagandist to upload new content, in contradiction with TikTok's own stated policy.

The
Washington
Post

Le Monde

B B C

THE
WALL STREET
JOURNAL



Aug 10th – Third Report

In August, we uncovered that content which appears to be banned could still be algorithmically promoted.

We call this unprecedented phenomenon *shadow-promotion*.

**The
Guardian**

euronews.

Le Monde

WIRED

Shadow-promotion: TikTok's algorithmic recommendation of banned content in Russia

Why is content that is supposed to have been banned since the start of the Ukraine-Russia war instead being promoted?

August 10 2022

reports@tracking.exposed



Impact

- Raising awareness with media coverage
- TikTok's CEO summoned by 6 US senators in an official letter based on our findings.
- TikTok was forced to clarify their policy, and fix its implementation

United States Senate
WASHINGTON, DC 20510

June 17, 2022

Mr. Shou Zi Chew
Chief Executive Officer
TikTok
5800 Bristol Parkway, Suite 100
Culver City, CA 90230

Dear Mr. Chew,

We write to request information about actions taken by TikTok to remove content or prohibit viewing of content relating to the ongoing war in Ukraine. Not only has TikTok banned users in Russia from livestreaming or posting new content, it has also blocked users in Russia from seeing any content uploaded by non-Russian TikTok accounts, creating an information blackout that deprives the Russian people access to objective reporting about the atrocities being committed in Ukraine.¹ Alarming, recent reports indicate TikTok's non-enforcement of its own purported moderation policies has allowed Russian state media to flood the platform with dangerous pro-war propaganda. No company should find itself in the position of amplifying the Kremlin's lies, which fuel public support for Russia's war of choice in Ukraine.²

On March 6, 2022, TikTok suspended livestreaming and new content on its video service in Russia to protect itself from a "fake news" law, which banned content critical of Russia's actions in Ukraine.³ We are deeply concerned that, through its failure to equally enforce this policy, TikTok is enabling the spread of pro-war propaganda to the Russian public, which risks adding to an already devastating human toll for both Ukrainians and Russians.

According to reporting by *The Washington Post*,⁴ TikTok's suspension of new video uploads and live streams from Russia has not been equally enforced. Research from Tracking Exposed shows that before TikTok's March 6 suspension, new uploads in Russia related to the conflict were roughly balanced between pro-war and anti-war content, but between March 6 and March 25, more than 93 percent were pro-war.⁵ Alarming, a number of content producers evading TikTok's suspension are Russian state-controlled media such as RIA Novosti, effectively making TikTok a conduit of state propaganda.⁶ While *The Washington Post* reports that it appears TikTok belatedly closed this loophole on March 25, the misleading, pro-regime

Adversarial audits will remain necessary

- Different observation scope
- No restriction on research question
- Guarantees data integrity



Join the effort!



- Development** - Put resources in common free-software repositories
- Policy** - Guarantee legal shield for public interest auditing
- Funding** - We need support to sustain our growth and hold platforms to account
- Monitoring** - Run your own audits, and contribute to the data safe

Get in touch: marc@tracking.exposed



[@trackingexposed](https://twitter.com/trackingexposed)
[@marcfaddoul](https://twitter.com/marcfaddoul)