

TikTok. Who's there?

A TikTok Algorithmic Audit Landscape

AI FORENSICS

Marc Faddoul

EU Disinfo Conference 2023

- Marc Faddoul
- Data Scientist and Transdisciplinary technologist
- Co-founder and Director of AI Forensics



TikTok Commission, French Senat



AI Forensics is a European non-profit which investigates influential and opaque algorithms.

AIF FORENSICS

Looking behind the code



So...
What's
so
special
about
TikTok?



INVISIBLE CENSORSHIP

TikTok Told Moderators to Suppress Posts by “Ugly” People and the Poor to Attract New Users



TikTok's Secret 'Heating' Button Can Make Anyone Go Viral

Forbes

Emily Baker-White
Jan 20, 2023

TikTok and ByteDance employees regularly engage in "heating," a manual push that ensures specific videos "achieve a certain number of video views," according to six sources and documents reviewed by Forbes.



It's not
just TikTok!

Elon Musk
@elonmusk

...



Elon's tweets

Twitter

640 x

6:35 AM · 14 févr. 2023 · 176,4 M vues

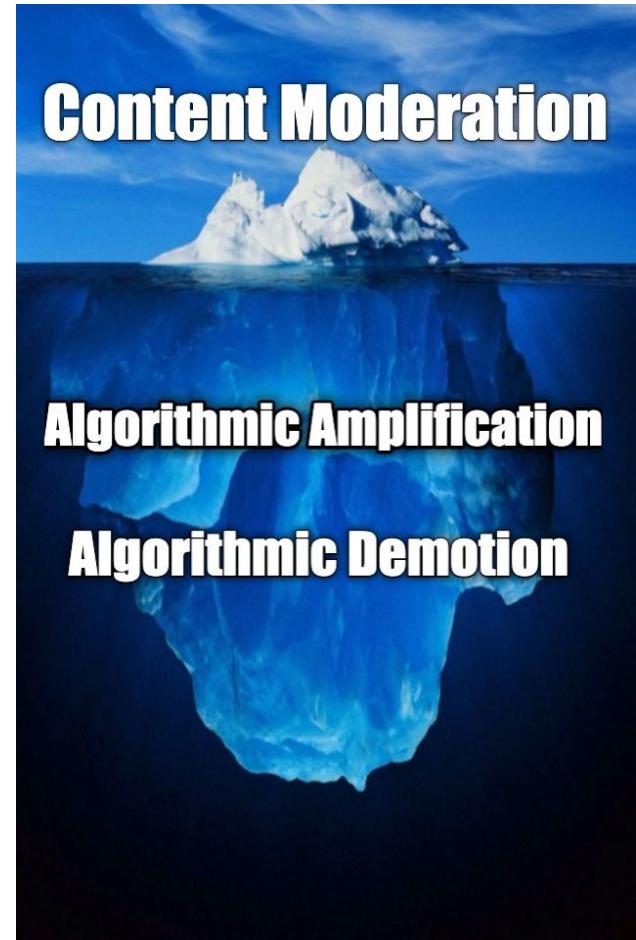
83,6 k Retweets 20,6 k citations 1,4 M J'aime 28,7 k Signets

Commenter Partager Aimer Marquer Partager Don

Why do we need algorithmic audits?

Recommendation algorithms are:

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



DSA acknowledges the issue. Does it solve it?

- Risk Assessment Reports (Art .34)
- Third-party Audits (Art. 37)
- Researcher Data access (Art. 40)
- Code of Practice engagements



First Tests: TikTok Ad Library

- Infringing content is deleted
SNV, 2023
- Missing sponsored content
Mozilla, 2021
- Access arbitrarily suspended
Marcus Bosch, 2023

TikTok

Ad Library All ads report Other commercial content FAQs

Find ads on TikTok

Ad target country: Poland Ad type: All ad types Ad published date: 10/1/2022–10/7/2023 Advertiser name or keyword: tusk Search

Search results

Total ads: 69 Last shown date: Newest to oldest

The screenshot shows the TikTok Ad Library interface. At the top, there are filters for 'Ad target country' (Poland), 'Ad type' (All ad types), 'Ad published date' (10/1/2022–10/7/2023), and 'Advertiser name or keyword' (tusk). Below the filters, the heading 'Search results' is displayed, followed by the message 'Total ads: 69' and 'Last shown date: Newest to oldest'. The results are presented in a grid of four cards. The first card is dark with a play button and the text 'wina Tuska'. The second card features a logo for 'DJeS Poland' and the text 'całość na moim kanale YouTube'. The third and fourth cards both feature the 'tusk:' logo and the text 'Od 2024 r. wprowadzimy LGBT do szkół podstawowych !!!'. The bottom half of these cards has a red background with white text.

TikTok Researcher API Terms of Service

- Submit project for approval
- Checking and removing deleted content
- *Worldwide, free, perpetual rights to research, and to researcher's institution brand*
- Quite limited quotas
- Academic affiliation



Content Collection VS Algorithmic Analysis



What data is available?

With the Research API, researchers can access public data on:

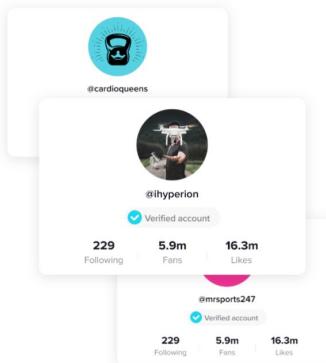


Accounts

This includes user profiles, and other data such as the number of likes of creator's content, followers and following count.



Content



For more information on the data returned via our APIs, check out the [Research API Codebook](#)

Key Missing Piece: Algorithmic Dynamics



User Engagement VS Algorithmic Promotion

Cooperative Audits

- Relies on official APIs and vetted access
- Various pain points and limitations
- Platform initiative or imposed by DSA

Adversarial Audits

- Data collected without platform consent
- Labor and expertise intensive
- Main driver of algorithmic accountability



Auditing TikTok's Algorithm: Let's get concrete

Individual Personalization Dynamics

- Create Account
- Define persona (profile, interests)
- Define user behavior
- Automate feed interaction
- Collect recommended content
- Analyse data



Image Credits: Junkipedia

TikTok is different on mobile and web!



Junkipedia : Issues

junkipedia.org

Junkipedia

Search Issues

Submit a Tip

Browse Issues

Issue Map

Issue Stats

MANAGE

Received Tips 242

My Issues 45

Manage Organization

Manage User Account

Manage Network

SAVED SEARCHES

Coronavirus Issues

Census Issues

Cures

Issues about Cures

Filter Issues

View As:

Sort By:

480 Issues

Coronavirus Treatment: Experimental Drug Shows 'Impressive' Results As COVID-19 Cure

International Business Times

Coronavirus Treatment: Experimental Drug Shows 'Impressive' Results As COVID-19 Cure

Most Read

How To Help During Covid-19

94

vaccine coronavirus cure

From Vidya Pradhan (via ncoc whatsapp tipline) 7 hours ago <https://www.ibtimes.com/coronavirus-treatment-experimental-drug-shows-impressive-results-as-covid-19-cure>

964-Issue

This Facebook post in "Michiganders Against Excessive Quarantine," states that the content posted is from Bill O'Riley who allegedly claims that COVID-19 is a bioweapon

Tig James Esq #FBPE #NotmyPM #Econocide on Twitter: "I live in Portugal. I am not rebounding from lies, lack of information, fudged death toll figures or ill-formed ideas as to the contingency plans. Every day we know how many cases, where and the death toll in each area. We have testing stations. Still think the UK is doing ok?"

1.000 Retweets 2.7K Likes

New to Twitter? Sign up Log in

Relevant people

The James Esq #FBPE #NotmyPM #Econocide

Trends for you

Minnie

Harper Lee

... view full issue

Coordination Malicious Framing death toll coronavirus

United Kingdom

Junkipedia's research interface

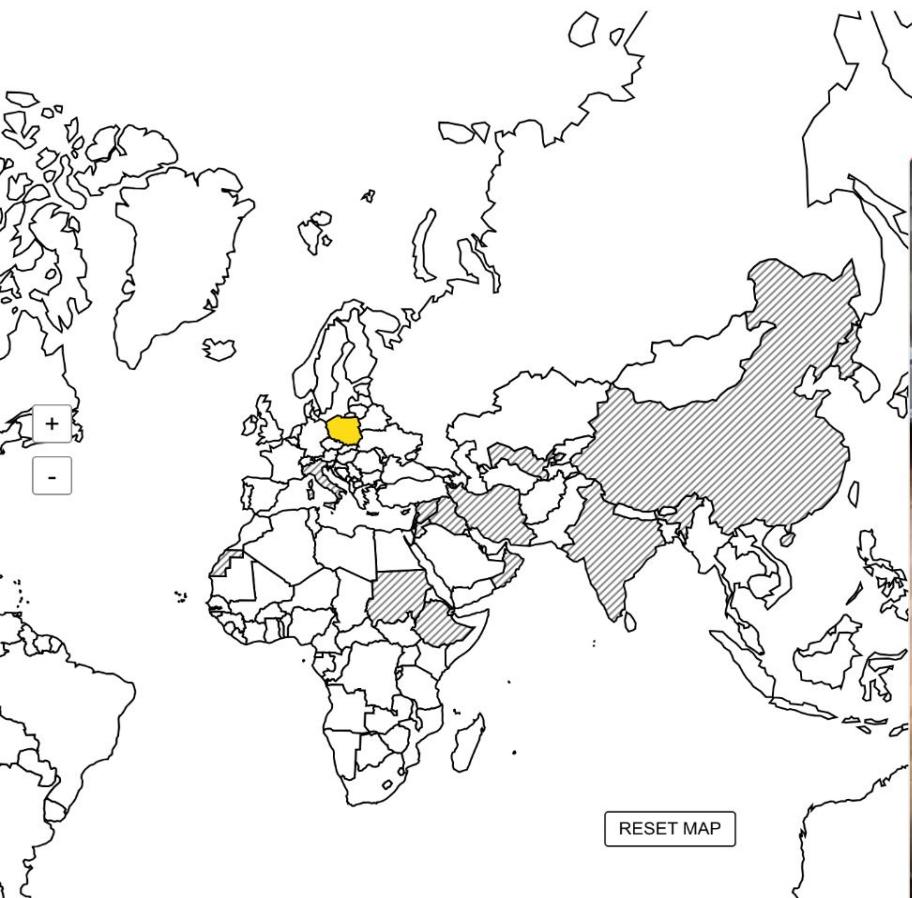
CROSS OVER

is an advanced auditing instrument that utilises in-home devices for genuine user simulation, providing a **detailed analysis of recommendation algorithms behaviors** and their contribution to disinformation.



CHOOSE TIMEFRAME

8/26/2023 – 9/9/2023



RESET MAP



Most Recommended Videos in Poland

26/08/2023 - 09/09/2023

[Download CSV](#)**Author Name:** stendzboco**Created Time:** 30/07/2023 18:35:44**Description:** #wiatrak**Music Title:** original sound**Author Name:** liverpoolfc**Created Time:** 31/07/2023 19:05:52**Description:** Your new Reds captain... Virgil van Dijk 🙌 #Liverpool #LFC #vandijk**Music Title:** LOOK AT ME**Author Name:** zdravievam**Created Time:** 05/06/2023 15:13:19**Description:** ПРИ БООЛЯХ В СУСТАВАХ И ТЯЖЕСТИ В НОГАХ ЭТА СМЕСЬ ЗАМЕНИТ ДОРОГИЕ МАЗИ И ГЕЛИ spine decompress**Music Title:** оригинальный звук - Здоровье**Author Name:** zdorovie2**Created Time:** 11/08/2023 00:48:59**Description:** @zdorov02 гречка с кефиром знаете почему эта диета не работает все потому что надо взять... Vegetables Drink**Music Title:** оригинальный звук - Здоровье и



Algorithmically driven Risk Scenarios

Auditing Recommender Systems

Putting the DSA into practice with
a risk-scenario-based approach



Search Auto-complete

5:11

LTE

- Q boys Cancel
- Q boys kissing girls buts ↵
- Q boys only click on the sound ↵
- Q boys.original  ↵
- Q boys kissing boys ↵
- Q boys snaps meaning ↵
- Q boys look at my liked videos ↵
- Q boys only dirty ↵
- Q boys wanna play with girls ↵
- Q boys will be bugs ↵
- Q boys click on this sound ↵

5:10

LTE

- Q girls Cancel
- Q girls tweekering on boys wieners ↵
- Q girls showing their privates ↵
- Q girls putting ice in their cooch ↵
- Q girls kissing girls in bed ↵
- Q girls in bikini ↵
- Q girls kissing girls with togue ↵
- Q girls tweekering hard ↵
- Q girls  ↵
- Q girlskissinggirls ↵
- Q girls bobbies ↵

Disinformation and propaganda



TikTok and Meta warned over Slovakia election lies

Elections: TikTok is political

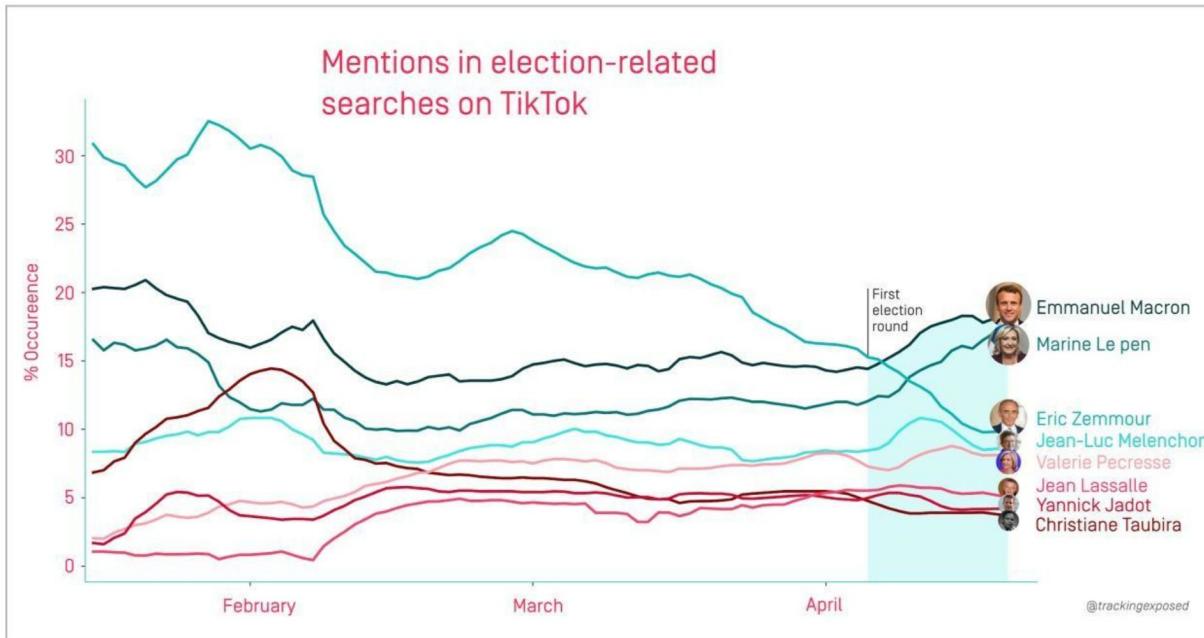
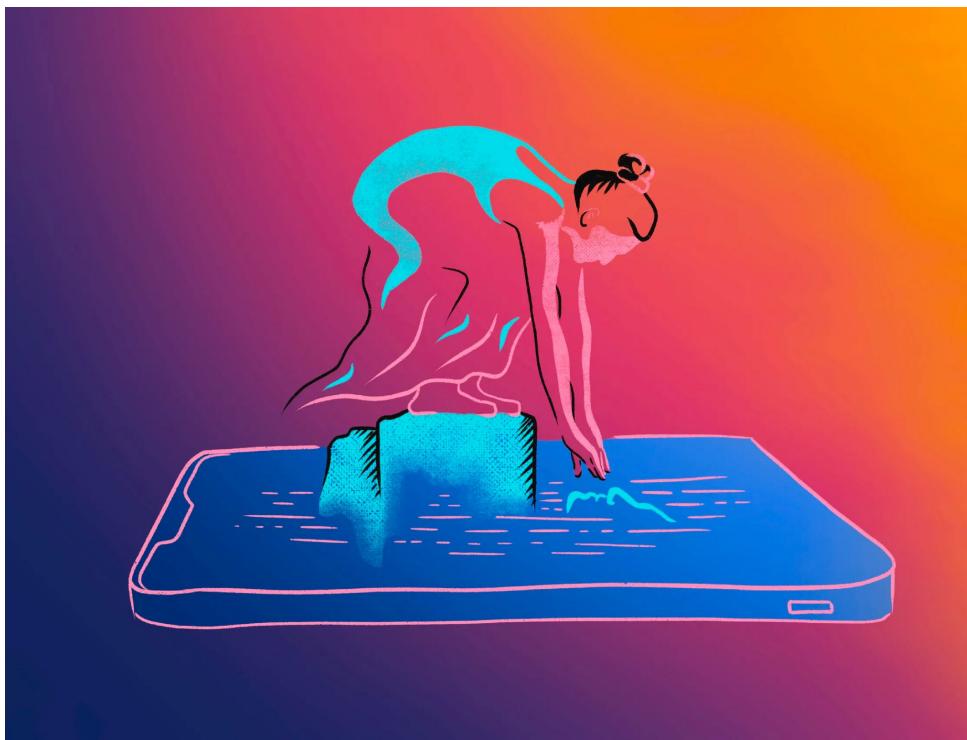
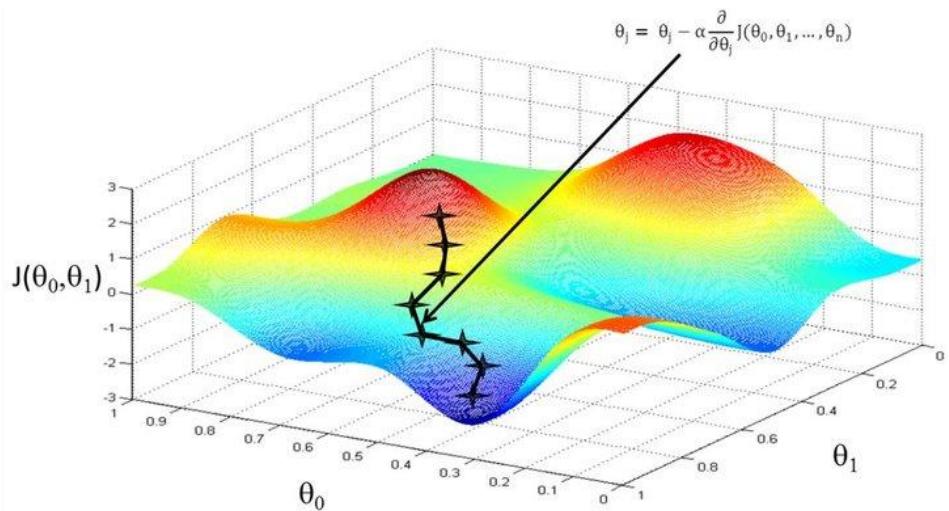


Figure 5: Prevalence of candidate terms in neutral election searches on TikTok

Mental Health: addictive self-harming content



Engagement VS User Value



What is "Value" for TikTok's algorithm

$$\text{Value} = P_{\text{like}} \times V_{\text{like}} + P_{\text{comment}} \times V_{\text{comment}} + E_{\text{playtime}} \times V_{\text{playtime}}$$

Value Model is for assigning value to interactions. Interaction values are divided into the following:

- 1. User value.** The value that the user who saw the video creates. Mostly includes playtime, retention and satisfaction etc.
- 2. Author value.** The value generated for authors, such as traffic, interactions, revenues etc.
- 3. Platform value.** The value created for the platform. Includes brand image, content safety, platform revenue etc.
- 4. Indirect value.** The value created indirectly by the video view, such as push notifications sent to users after they have left comments, might improve retention and affect the content ecosystem, etc

BRIEFING NOTE

Fixing Recommender Systems

From identification of risk factors to
meaningful transparency and mitigation



Conclusion

AIF FORENSICS

Looking behind the code

Marc Faddoul
marc@aiforensics.org

AIF



COMMISSION D'ENQUÊTE **INFLUENCE TIKTOK**



Le fait que l'application reste dans l'orbite des autorités chinoises pourrait permettre à celles-ci, par exemple, de **reconstituer, via les données collectées, l'organigramme** d'une organisation, administration ou entreprise, afin d'en identifier les éléments pertinents en vue **d'initier des actions d'espionnage ; d'identifier les habitudes** de ces personnes pour faciliter ces mêmes actions ; de collecter des informations précises sur une personne pour construire un e-mail personnalisé pertinent dans le cadre d'une **cyberattaque** ; d'identifier **certaines lieux particulièrement stratégiques pour l'État** et les personnes qui les fréquentent de manière habituelle ; de lancer **une campagne de désinformation** lors d'une crise internationale au profit de la Chine ou de ses alliés ; de **favoriser, lors d'élections nationales ou locales, certains candidats.**

How it works



- Raspberry Pi computers
- LineageOS
- MITM Proxy
- Orchestrator Application
- Server Application