

# IS TIKTOK A SYSTEMIC RISK FOR DEMOCRACY?

THE DIGITAL SERVICES ACT AND NEW RULES FOR  
ALGORITHMIC ACCOUNTABILITY

Dr. Martin Degeling

AI Forensics

13.12.2024

# ABOUT ME

- Research on **Usable Privacy and Security, the GDPR and Web Privacy** (Ruhr University, CMU)
- since 2022: Research at **Civil Society Organisations**: Stiftung Neue Verantwortung, Institute for Strategic Dialoge, AI Forensics

AI FORENSICS

# SUMMARY

- Intro to the **Digital Services Act (DSA)**
- Basics about **Recommender Systems**
- Insights on recent studies about TikTok and **elections**

Slides at: [martin.degeling.com/slides/1312](https://martin.degeling.com/slides/1312)

# INTRODUCTION TO THE DSA

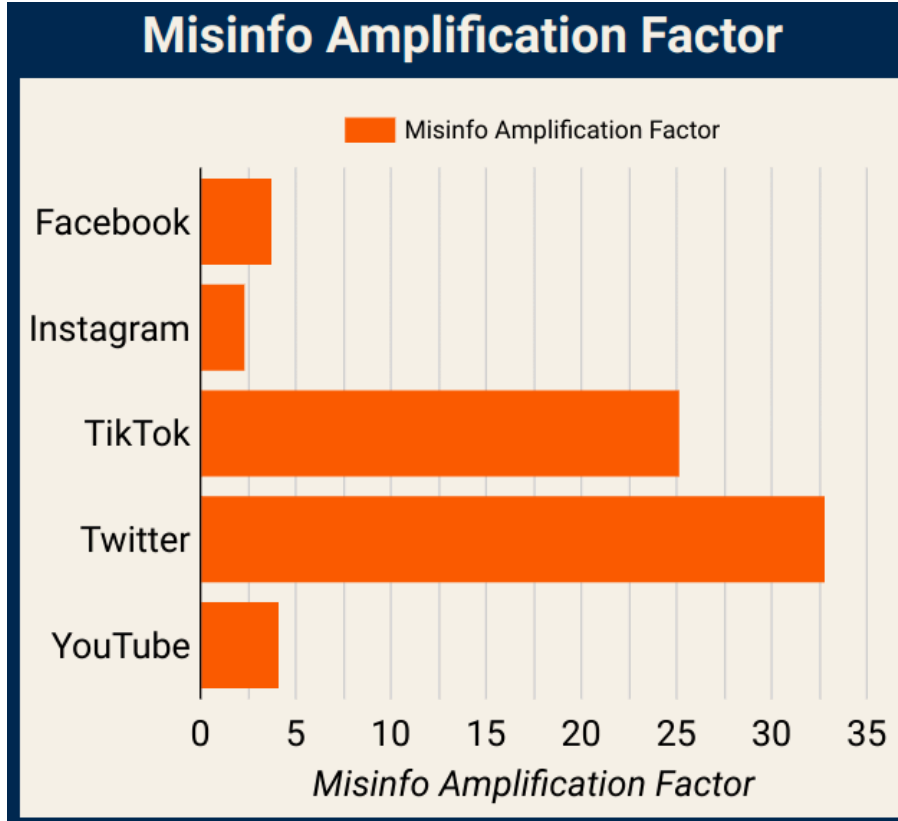


"The **Digital Services Act** is a EU Regulation that defines obligations for online services regarding liability for illegal content, content moderation, transparency and due diligence obligations for service providers."

Source: EU Commission



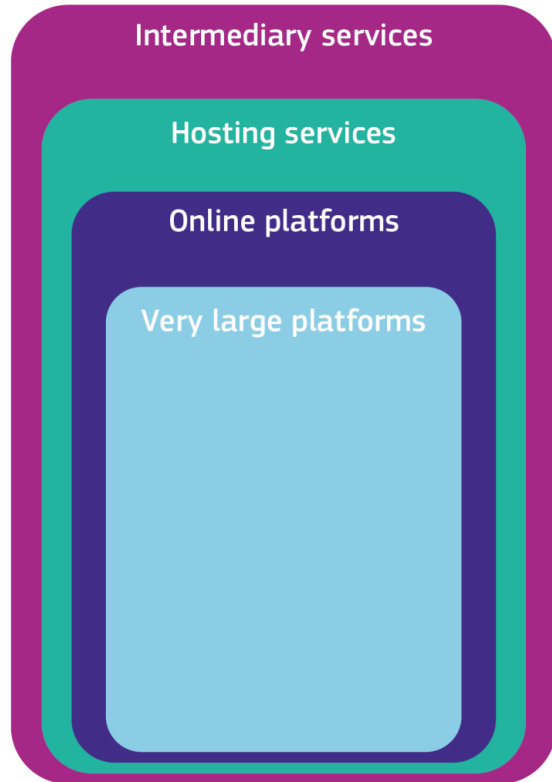
# WHY THE DSA?



Integrity Institute, 02/2023

- Cambridge Analytica and fears around **micro-targeting**
- **Filter Bubbles** and Echo Chambers (*although debunked*)
- **Dis-** and **Misinformation** (*around Elections and Covid*)
- **Hate Speech** Online

# WHO IS REGULATED BY THE DSA?



- **intermediaries:** online service that only passes on data
- **hosting:** services that stores information
- **online platform:** a hoster that disseminates user information to the public
- **VLOPs and VLOSEs:** online platforms with more than 45 M monthly users in the EU

# WHAT ARE THE NEW RULES FOR WHOM?

New obligations	Intermediary services	Hosting services	Online platforms	Very large platforms
Transparency reporting	•	•	•	•
Requirements on ToS due account of fundamental rights	•	•	•	•
Cooperation with national authorities	•	•	•	•
Points of contact	•	•	•	•
Notice and action/provide information to users		•	•	•
Reporting criminal offenses		•	•	•
Complaint and redress mechanism and out of court dispute settlement			•	•
Trusted flaggers			•	•
Measures against abusive notices and counter-notices			•	•

# WHAT ARE THE NEW RULES FOR WHOM?

New obligations	Intermediary services	Hosting services	Online platforms	Very large platforms
Special obligations for marketplaces			•	•
Bans on targeted adverts to children			•	•
Transparency of recommender systems			•	•
User-facing transparency of online advertising			•	•
Risk management obligations and crisis response				•
External & independent auditing				•
User choice for recommender				•
Data sharing with authorities and researchers				•
Codes of conduct				•
Crisis response cooperation				•

# WHAT ARE THE VLOPS/VLOSES?

Very Large Online  
Platforms/Search Engines  
with more than 45 Million  
monthly active users

- **Social Media:** Facebook, Instagram, LinkedIn, Pinterest, Snapchat, TikTok, X, Youtube
- **Shopping:** Alibaba AliExpress, Amazon, Google Shopping, Temu, Zalando
- **Search:** Google, Bing
- **Adult Content:** Pornhub, Stripchat, XNXX, XVideos
- **Other:** Apple AppStore, Booking.com, Google Maps, Google Play, Wikipedia

# WHY? BECAUSE SELF-REGULATION DIDN'T WORK

Many of the measures existed before

- There already was a **code of practice on misinformation** in the EU
- Facebook offered **Crowdtangle**, a tool for researchers to study content distribution
- Some **Transparency** measures exist (some thanks to GDPR)
- **Trusted flaggers** (Youtube heroes) already exist

# WHY? BECAUSE SELF-REGULATION DIDN'T WORK

Many of the measures existed before

- There already was a **code of practice on misinformation** in the EU, *but not required so Twitter withdrew*
- Facebook offered **Crowdtangle**, a tool for researchers to study content distribution
- Some **Transparency** measures exist (some thanks to GDPR)
- **Trusted flaggers** (Youtube heroes) already exist

# WHY? BECAUSE SELF-REGULATION DIDN'T WORK

Many of the measures existed before

- There already was a **code of practice on misinformation** in the EU, *but not required so Twitter withdrew*
- Facebook offered **Crowdtangle**, a tool for researchers to study content distribution, *but it's been killed and no other platform has something similar*
- Some **Transparency** measures exist (some thanks to GDPR)
- **Trusted flaggers** (Youtube heroes) already exist



# WHY? BECAUSE SELF-REGULATION DIDN'T WORK

Many of the measures existed before

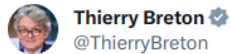
- There already was a **code of practice on misinformation** in the EU, *but not required so Twitter withdrew*
- Facebook offered **Crowdtangle**, a tool for researchers to study content distribution, *but it's been killed and no other platform has something similar*
- Some **Transparency** measures exist (some thanks to GDPR), *but others like the Twitter & Reddit API were closed down*
- **Trusted flaggers** (Youtube heroes) already exist

# WHY? BECAUSE SELF-REGULATION DIDN'T WORK

Many of the measures existed before

- There already was a **code of practice on misinformation** in the EU, *but not required so Twitter withdrew*
- Facebook offered **Crowdtangle**, a tool for researchers to study content distribution, *but it's been killed and no other platform has something similar*
- Some **Transparency** measures exist (some thanks to GDPR), *but others like the Twitter & Reddit API were closed down*
- **Trusted flaggers** (Youtube heroes) already exist, *but adoption was sparse*

# LEGAL REGULATION WAS NECESSARY..



Today @elonmusk and I wanted to share a quick message with you on platform regulation 🇪🇺

#DSA



11:16 PM · May 9, 2022 from Austin, TX

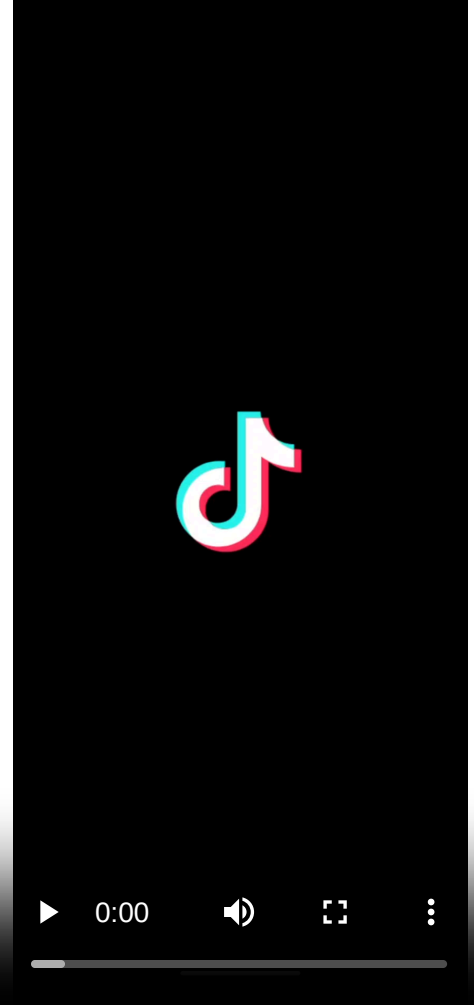
- .. to get from self-regulation to common standards
- .. raise the bar and level the playing field

# WHY ARE WE LOOKING AT TIKTOK?



- TikTok is a **designated VLOP** (> 150 Mio monthly users in the EU)
- "Relatively" **new and less studied**
- **immense impact** on it's users as well as the ecosystem

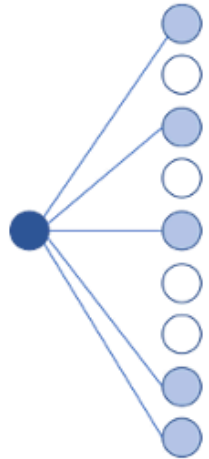
# A TIKTOK HANDS ON



- Shows **videos** in a constant stream (swipe up), allows editing
- Interaction through **likes and comments**
- Various **search** functions to explore content
- Financed through **ads, micro-payments**
- myriads of features (not shown: lives, shopping)

# WHAT WAS SO NEW ABOUT THE TIKTOK MODEL

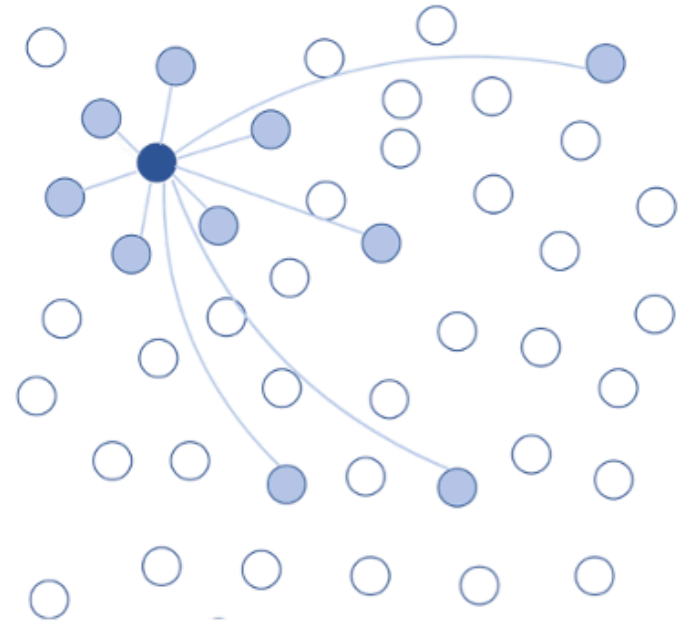
Subscriptions



Network



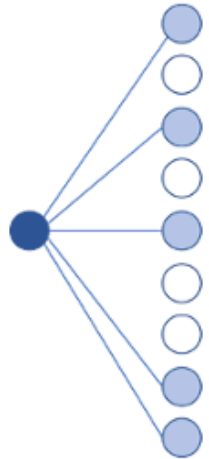
Algorithm



Graphic: Arvind Narayanan: Understanding Social Media Recommendation Algorithms, 3.9.2023

# WHAT WAS SO NEW ABOUT THE TIKTOK MODEL

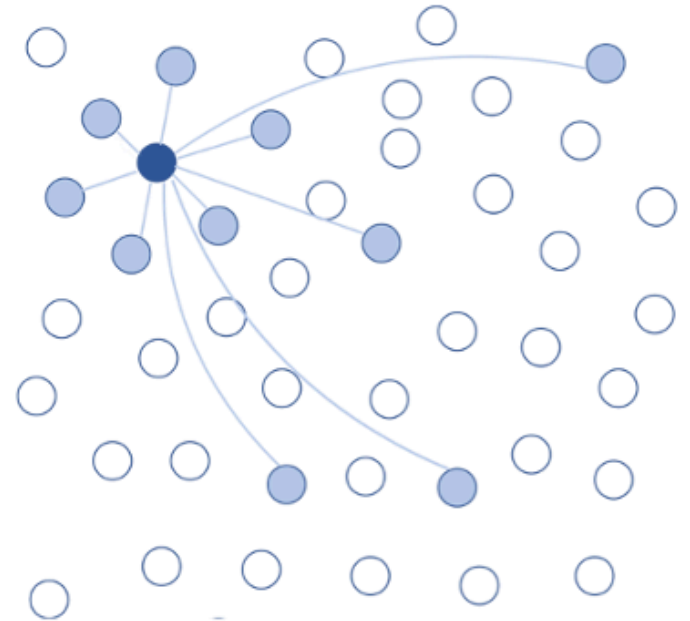
Subscriptions



Network



Algorithm





e.g. Podcasts, RSS

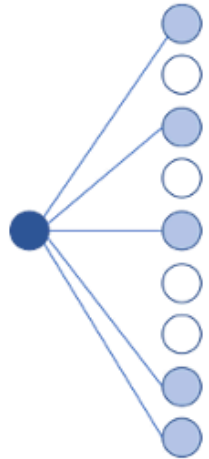
e.g. the "old" Facebook, IG

e.g. Tiktok FYP, IG Reels

Graphic: Arvind Narayanan: Understanding Social Media Recommendation Algorithms, 3.9.2023

# WHAT WAS SO NEW ABOUT THE TIKTOK MODEL

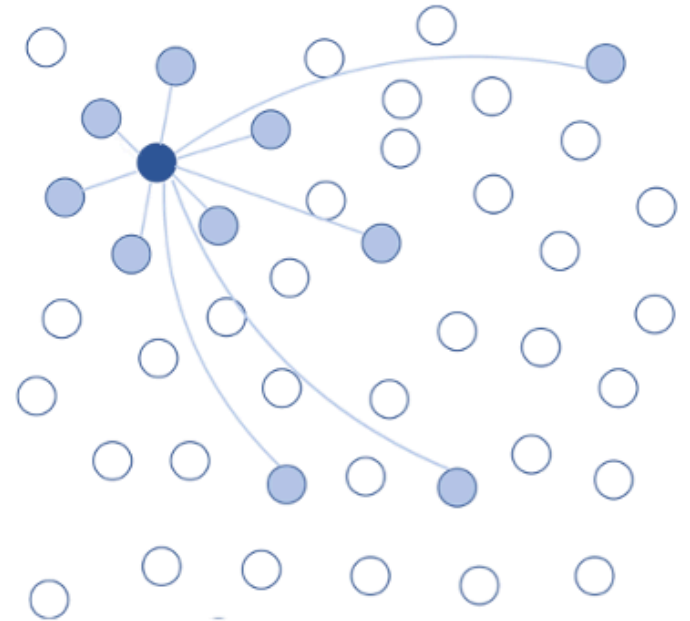
Subscriptions



Network



Algorithm



e.g. Podcasts, RSS

active selection

e.g. the "old" Facebook, IG

selection of others

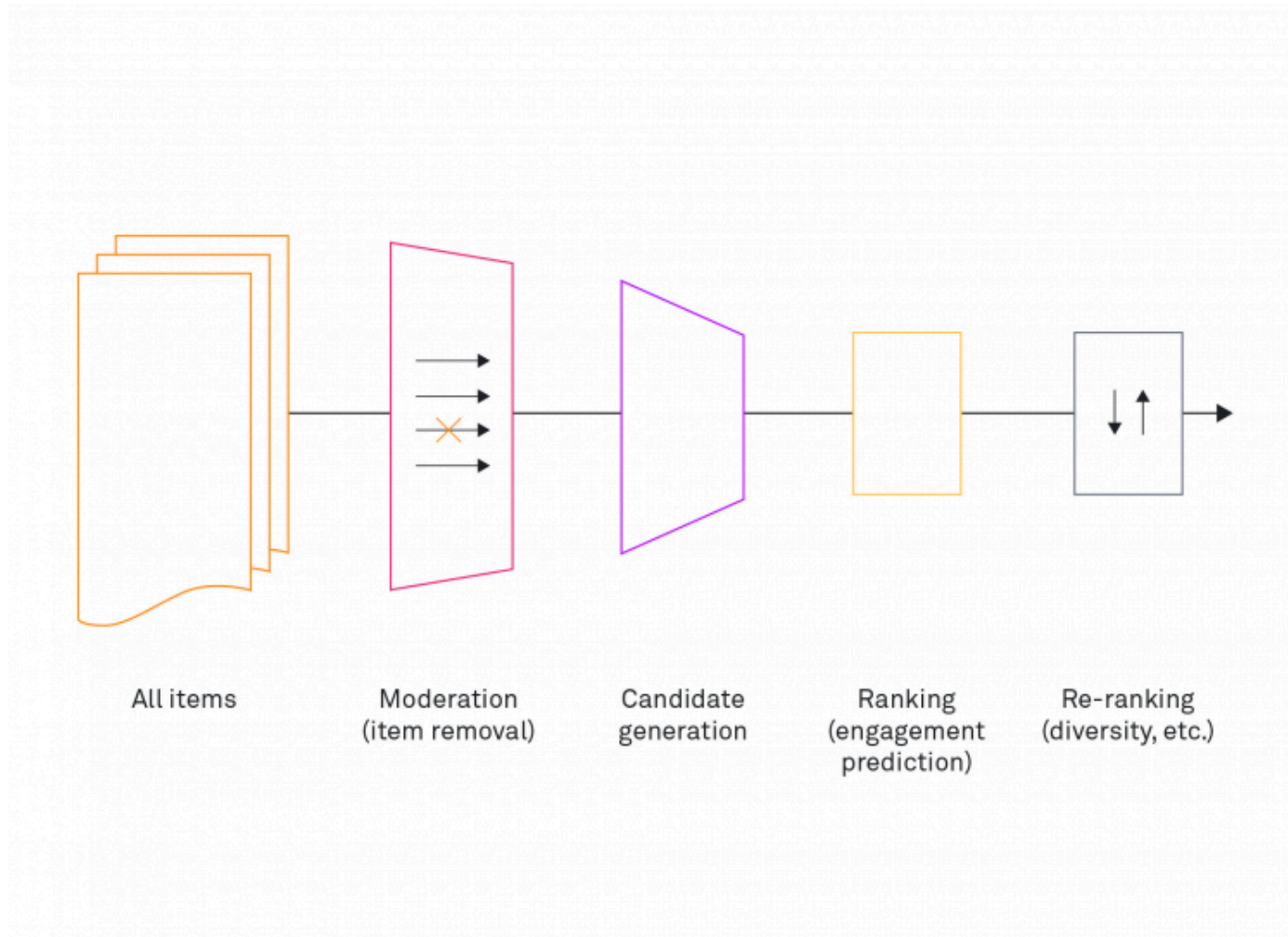
e.g. Tiktok FYP, IG Reels

weights (based on implicit feedback)

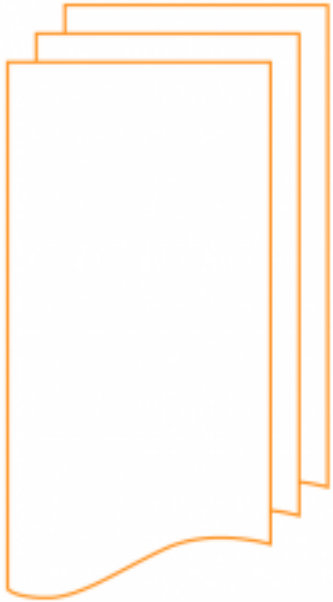
Graphic: Arvind Narayanan: Understanding Social Media Recommendation Algorithms, 3.9.2023

# A RECOMMENDATION ALGORITHM

# A RECOMMENDATION ALGORITHM



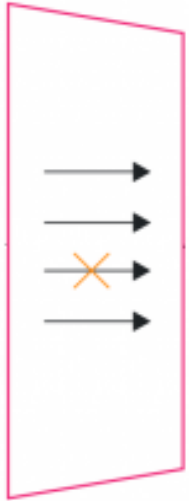
# RECOMMENDATION BASICS



All items

- **34 billion** videos are uploaded to TikTok daily
- The average watchtime for a video is **18 seconds**
- People watch TikTok **~58 minutes per day**.
- The recommender system has to select **190 out of 34.000.000.000** for you

# RECOMMENDATION BASICS

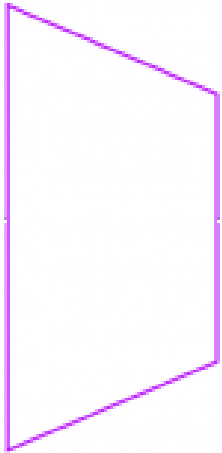


Moderation  
(item removal)

## MODERATION

- TikTok makes **3 Million** (automated) content moderation decision per day for Austrian content
- **1.7 Million videos** are removed for Austrian audience every day

# RECOMMENDATION BASICS



Candidate  
generation

## CANDIDATE GENERATION (OFFICIAL)

- **User interactions:** like, shares, comments..
- **Video information:** captions, sounds, hashtags.
- **Devices and account settings:** language, country, device. (less important)



# RECOMMENDATION BASICS

## CANDIDATE GENERATION (INTERNAL):

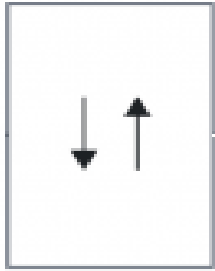
- Optimization on **usage time** and **user retention**
- Each video is assigned a **Value** for each user for how likely it will result in an interaction
- **Important:** Calculations are mostly based on metadata (who posted, how many likes are there already), less on the actual *content*



Ranking  
(engagement  
prediction)

Ben Smith, NYTimes: How TikTok Reads Your Mind, 6.12.2021, Translated Memo

# RECOMMENDATION BASICS



Re-ranking  
(diversity, etc.)

## RE-RANKING:

- The For You feed is based on sets of 8 videos
- The final re-ranking ensures that two **advertisements** are added.

# WHAT'S THE PROBLEM?

- it's not that TikTok is making young people **more stupid**, or China is using it to **destroy the western world**

Why TikTok Is a Threat to Democracy

SHARE  
Facebook  
Twitter  
LinkedIn  
Reddit  
Email



# WHAT'S THE PROBLEM?

- it's not that TikTok is making young people **more stupid**, or China is using it to **destroy the western world**
- Platforms optimize for **views and ad revenue**, not for media literacy or democratic values

Why TikTok Is a Threat to Democracy

SHARE  
•  
•  
•  
•  
•



# WHAT'S THE PROBLEM?

- it's not that TikTok is making young people **more stupid**, or China is using it to **destroy the western world**
- Platforms optimize for **views and ad revenue**, not for media literacy or democratic values
- Algos are prone to **bias and manipulation** (by humans)

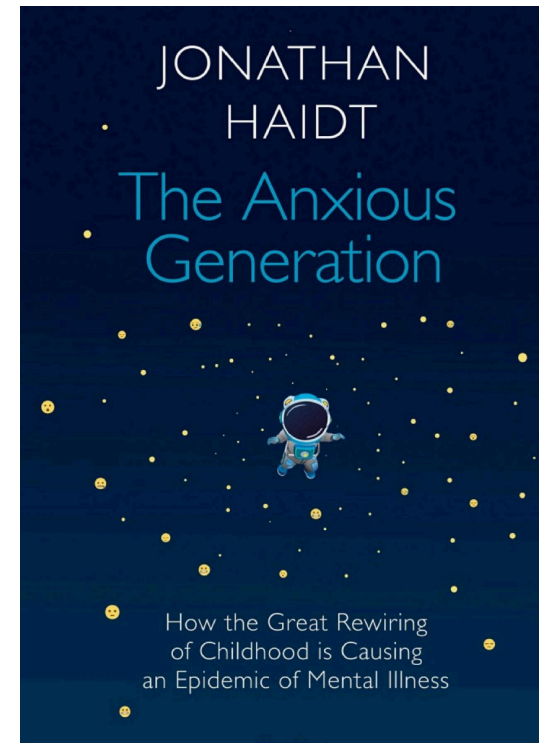
Why TikTok Is a Threat to Democracy

SHARE  
•  
•  
•  
•  
•



# WHAT'S THE PROBLEM?

- it's not that TikTok is making young people **more stupid**, or China is using it to **destroy the western world**
- Platforms optimize for **views and ad revenue**, not for media literacy or democratic values
- Algos are prone to **bias and manipulation** (by humans)



THE DSA ANSWER:

# RISK ASSESSMENTS

They are required by VLOPs to conduct on their own platform.(But we have little trust that they will be thorough so researcher and civil society should do them, too.)

# SYSTEMIC RISKS LISTED IN THE DSA (1/2)

- risks associated with the dissemination of **illegal content**, such as: child sexual abuse material, hate speech or other types of misuse of their services for illegal activities. *(Recital 80)*
- impact of the service **on the exercise of fundamental rights**, as protected by the EU Charter, including: human dignity, freedom of expression, media freedom and pluralism, the right to private life, data protection, the right to non-discrimination, the rights of the child and consumer protection. *(Recital 81)*



# SYSTEMIC RISKS LISTED IN THE DSA (2/2)

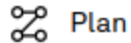
- the actual or foreseeable **negative effects on democratic processes, civic discourse and electoral processes, as well as public security.** *(Recital 82)*
- negative effect on the **protection of public health, minors and serious negative consequences to a person's physical and mental well-being, or on gender-based violence.** *(Recital 83)*

# SYSTEMIC RISKS LISTED IN THE DSA (2/2)

actual or foreseeable negative effects on democratic processes, civic discourse and electoral processes, as well as public security. *(Recital 82)*

- negative effect on the protection of public health, minors and serious negative consequences to a person's physical and mental well-being, or on gender-based violence. *(Recital 83)*

# FROM A VERY ABSTRACT RISK TO CONCRETE SCENARIO



Plan



Scenarios



Measure



Evaluate

Types of audits

Prioritise

Develop measurements to understand the scenario. There are different **types of algorithm audits**, as well as **platform elements** to consider. These can range from automated measurements that look at the actual implementation to user perspectives through surveys. An auditor needs to develop multiple measurements and then prioritise them to find the best measurement(s) to test a specific scenario.

[More details](#)

Select audit type or platform element:

Code audit



Connected elements

Algorithmic logic

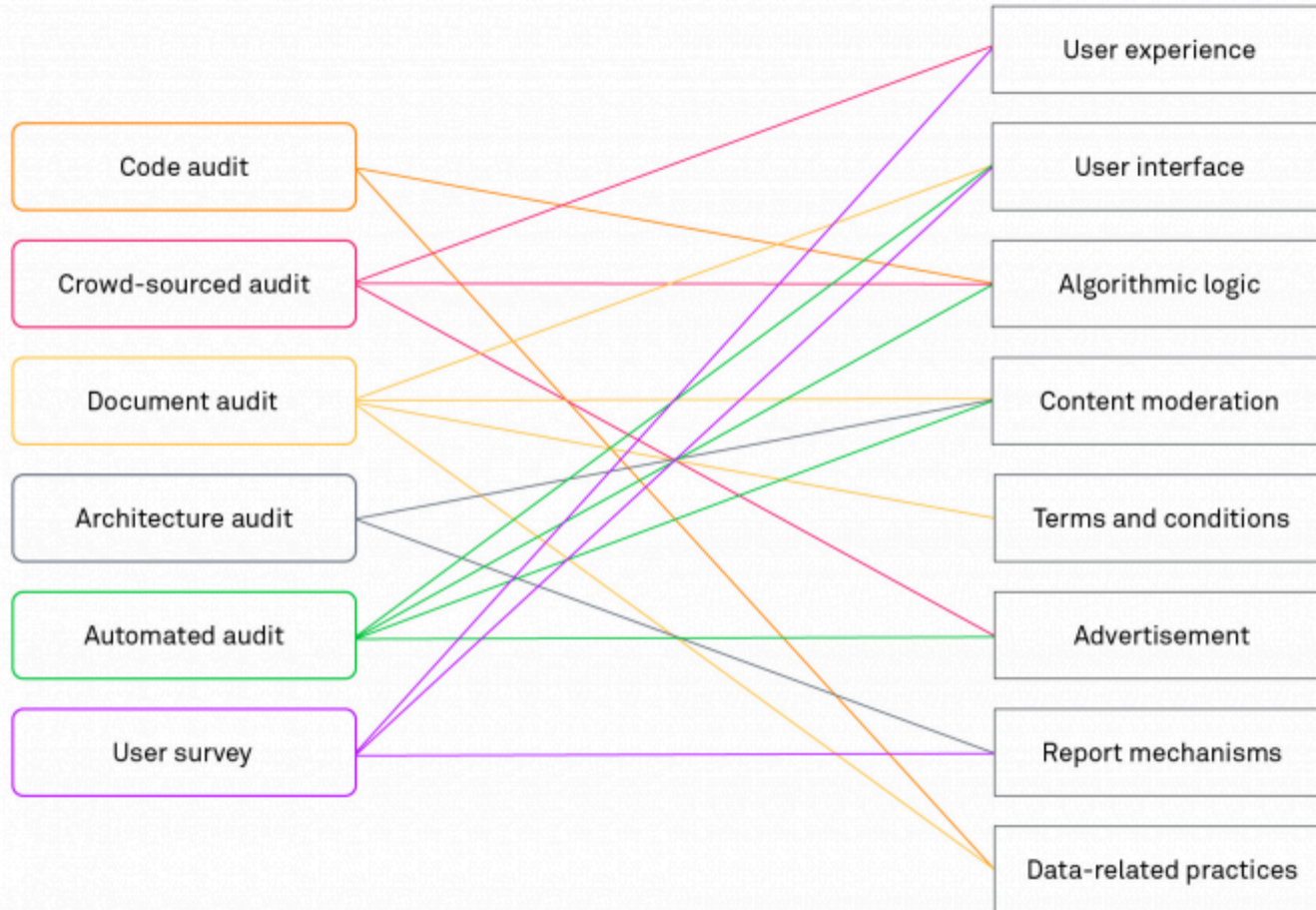
Data-related practices

Example:

Review source code,  
evaluate model parameters.



# HOW TO STUDY SYSTEMIC RISKS

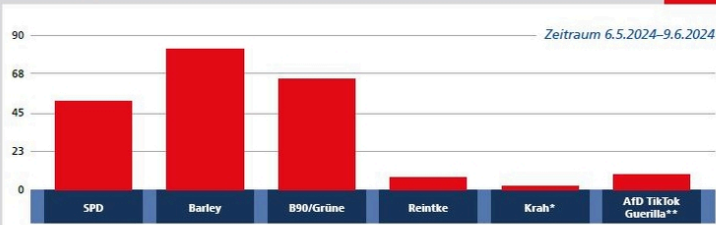


# EXAMPLES OF WHAT WE KNOW ABOUT TIKTOKS ALGORITHMS

# TIKTOK IN EU ELECTIONS

Anzahl der veröffentlichten TikTok-Videos der ausgewählten Accounts

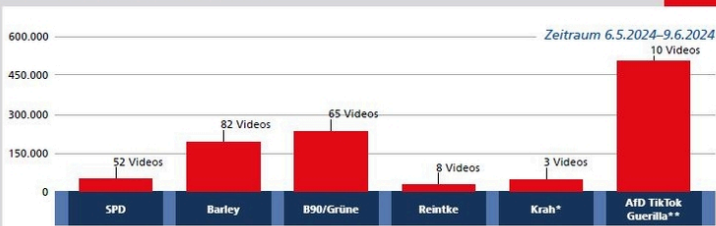
Abb. 1



\* Die Reichweite des offiziellen Accounts von Maximilian Krah wurde im Vorfeld der Europawahl von TikTok gedrosselt. Krah postete im genannten Zeitraum lediglich drei Videos und setzte stattdessen auf eine partizipative Verbreitung mittels „TikTok Guerilla“.  
\*\* Bis zum 9.6.2024 wurden insgesamt 5.248 Videos mit dem Hashtag #Krah gepostet. Ausgewertet wurden die Top Ten der im Berichtszeitraum geposteten Videos mit dem Hashtag #Krah.  
QUELLE: eigene Darstellung.

Anzahl der Views der veröffentlichten TikTok-Videos der ausgewählten Accounts

Abb. 2



\* Ausgewertet wurden die Top Ten der im Berichtszeitraum geposteten Videos mit dem Hashtag #Krah.  
QUELLE: eigene Darstellung.

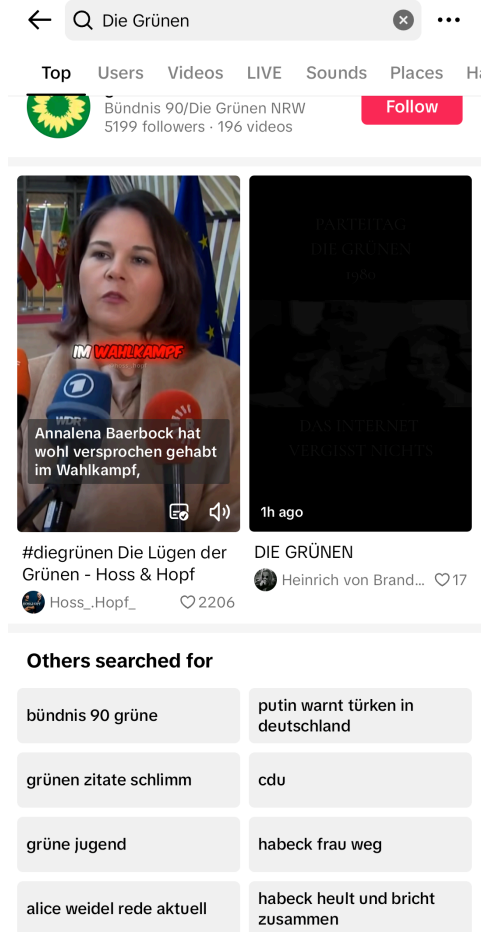
- **Research Question:** How was TikTok used in the EU election campaign in Germany?
- **Methodology:** Qualitative Analysis.
- **Findings:** AFD Guerrilla tactics, targeted audience approach

# SEARCH SUGGESTIONS

- **Research Question:** What does the search suggestion algorithm refer to
- **Methodology:** Quantitative & Qualitative Analysis.
- **Findings:** Conspiracies and references to AFD

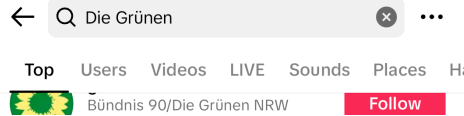


# SEARCH SUGGESTIONS

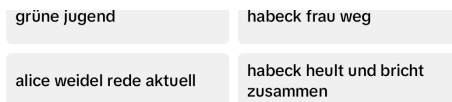
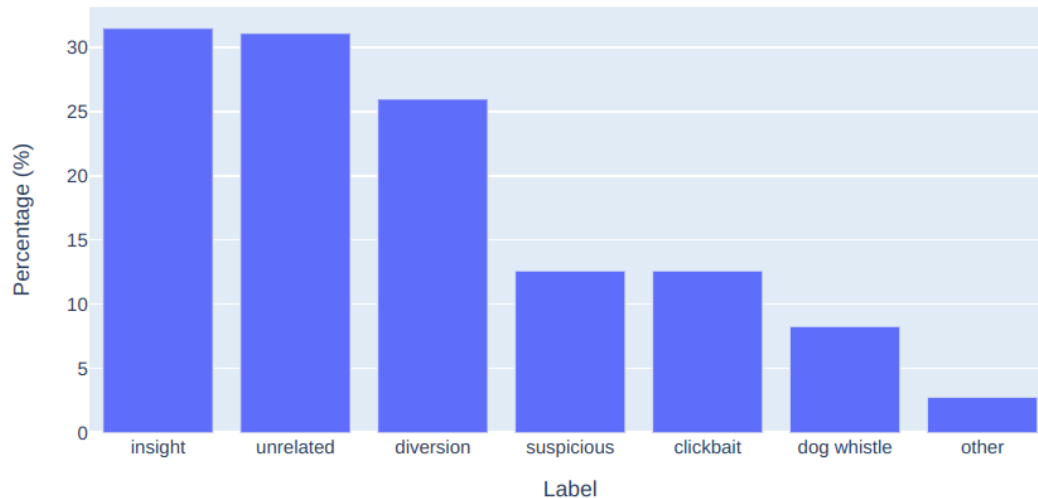


- **Research Question:** What does the search suggestion algorithm refer to
- **Methodology:** Quantitative & Qualitative Analysis.
- **Findings:** Conspiracies and references to AFD

# SEARCH SUGGESTIONS

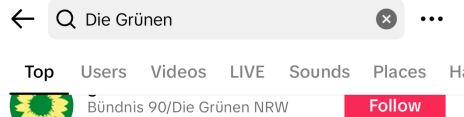


Occurrence of Labels in Search Suggestions for Political Parties

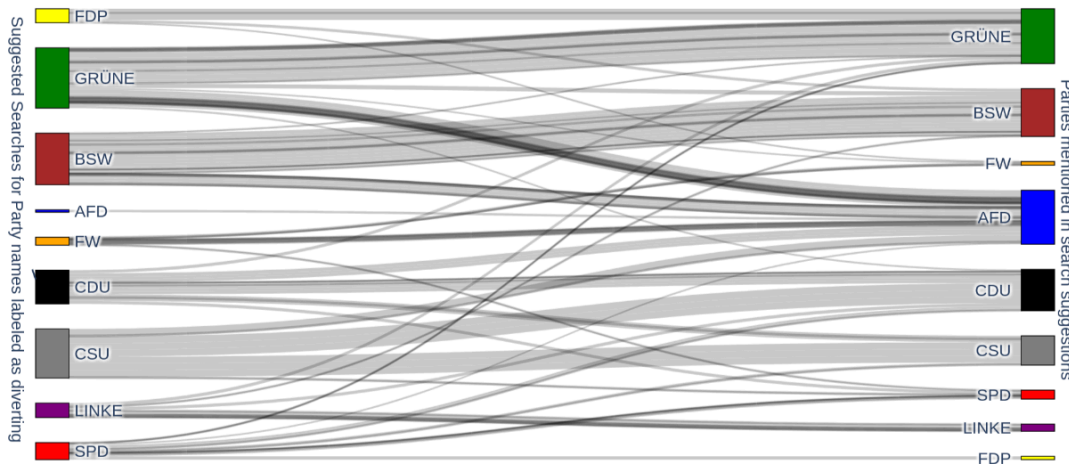


- **Research Question:** What does the search suggestion algorithm refer to
- **Methodology:** Quantitative & Qualitative Analysis.
- **Findings:** Conspiracies and references to AFD

# SEARCH SUGGESTIONS



Search Suggestions Mentioning Other Parties



grüne jugend      habeck frau weg

alice weidel rede aktuell      habeck heult und bricht zusammen

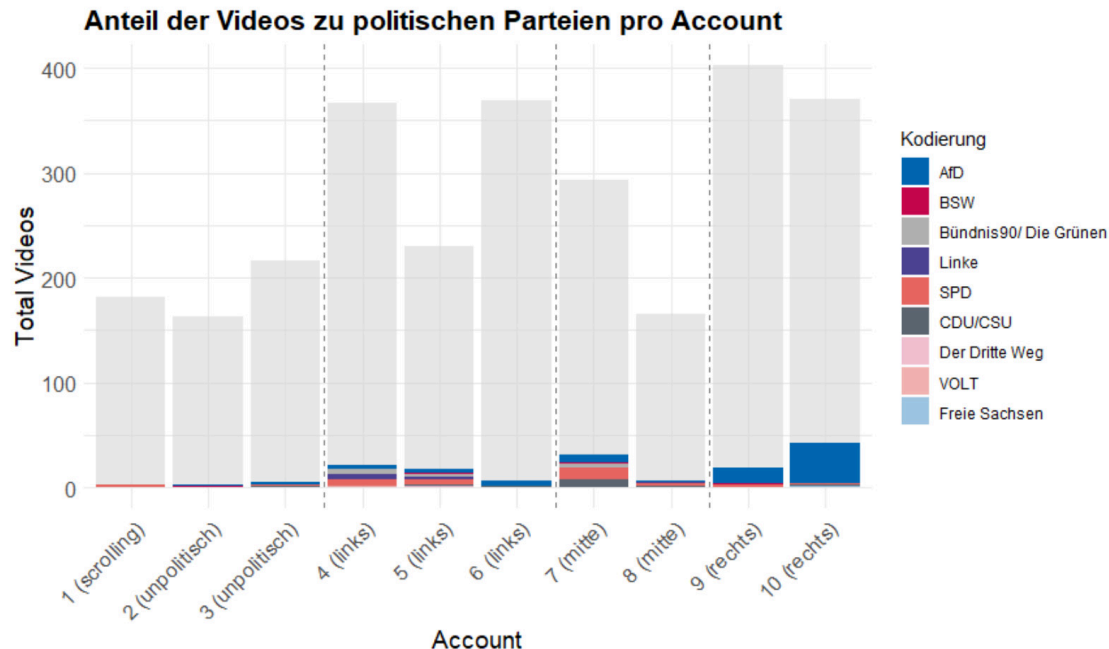
- **Research Question:** What does the search suggestion algorithm refer to
- **Methodology:** Quantitative & Qualitative Analysis.
- **Findings:** Conspiracies and references to AFD

Miazia Schüler, Martin Degeling, Salvatore Romano and Kathy Meßmer, Other searched for... , 09/2024

# FOR YOU IN BRANDENBURG

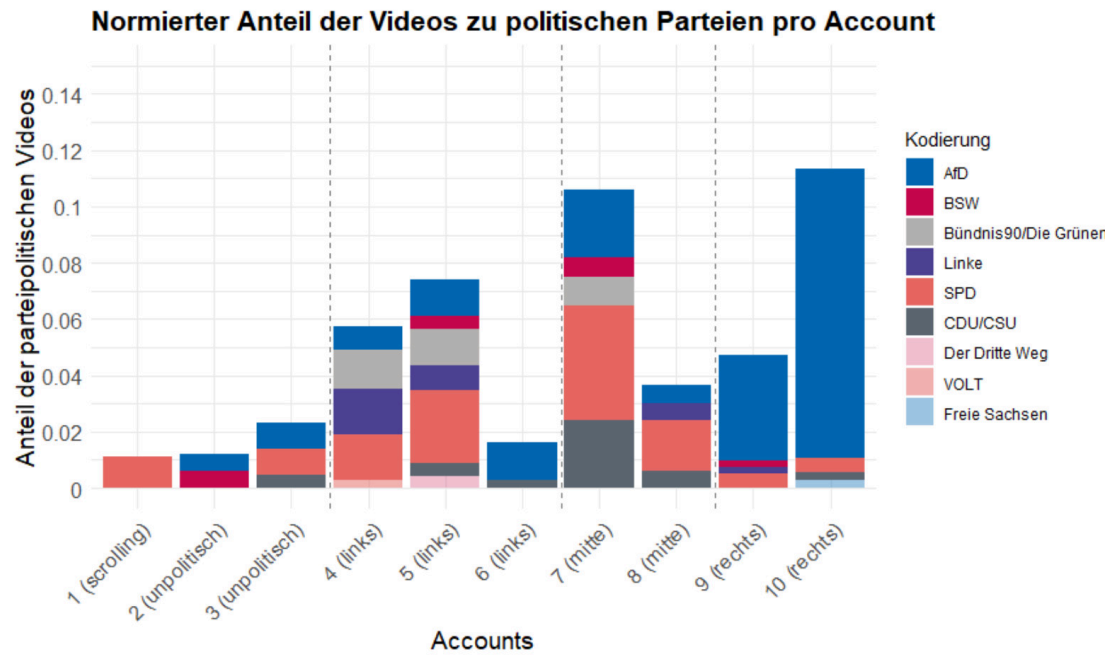
- **Research Question:** What content did personas of different political positions see during the election campaign in Brandenburg.
- **Methodology:** Qualitative and Quantitative Analysis.
- **Findings:** AFD is overrepresented, content is not recognized as political.

# FOR YOU IN BRANDENBURG



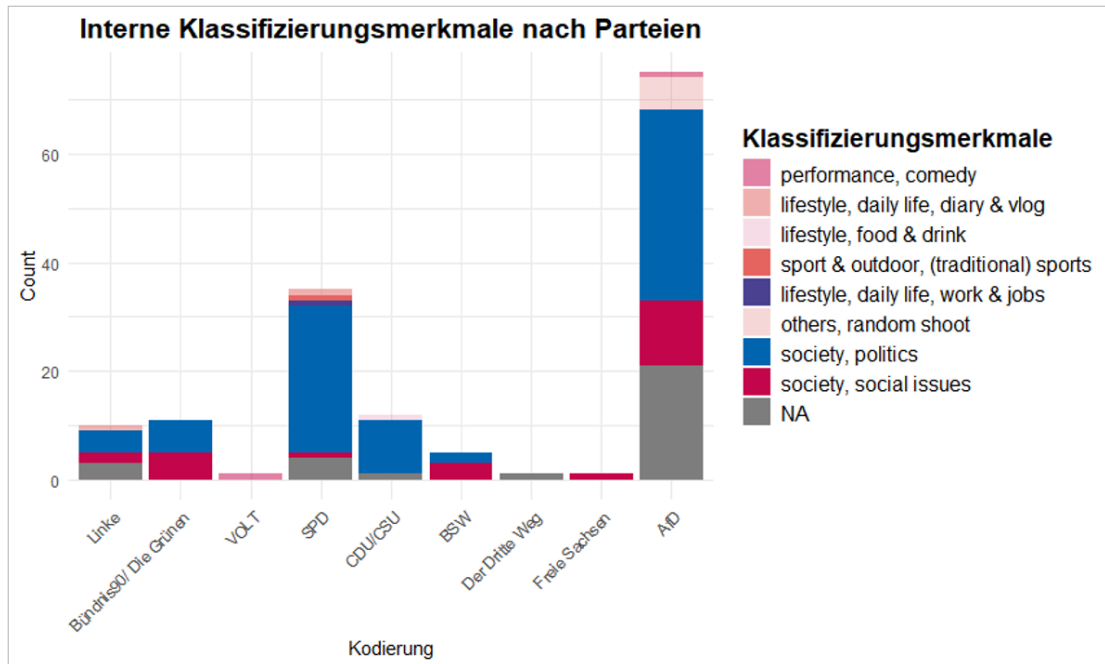
- **Research Question:** What content did personas of different political positions see during the election campaign in Brandenburg.
- **Methodology:** Qualitative and Quantitative Analysis.
- **Findings:** AfD is overrepresented, content is not recognized as political.

# FOR YOU IN BRANDENBURG



- **Research Question:** What content did personas of different political positions see during the election campaign in Brandenburg.
- **Methodology:** Qualitative and Quantitative Analysis.
- **Findings:** AfD is overrepresented, content is not recognized as political.

# FOR YOU IN BRANDENBURG



- **Research Question:** What content did personas of different political positions see during the election campaign in Brandenburg.
- **Methodology:** Qualitative and Quantitative Analysis.
- **Findings:** AFD is overrepresented, content is not recognized as political.

# WHAT ABOUT THE ROMANIAN ELECTION?





## UNGEWÖHNLICHER WAHLKAMPF

## Tiktok statt TV führte den Rechtsextremen Georgescu in Rumänien zum Wahlsieg

Der parteilose Kandidat für das rumänische Präsidentenamt verzichtete auf TV-Debatten oder Plakate und führte einen einseitigen Social-Media-Wahlkampf – mit Erfolg

27. November 2024, 05:00

Später lesen



Călin Georgescu hielt nach seinem Wahlsieg eine Pressekonferenz via Facebook ab.

EPA/ROBERT GHEMENT



ZEITUNG MEHR F.A.Z.

Abo



PLATTFORM GEGEN DEMOKRATIE

## Wie TikTok in Rumänien Wahlen mitentscheidet



Ein Kommentar von Michael Hanfeld

09.12.2024, 18:36 Lesezeit: 2 Min.



# WHAT WE KNOW



- Georgescu focused on TikTok for **marketing**
- Several (small) **coordinated influence campaigns** were stopped by TikTok
- 100,000 fake likes and followers were **deleted**
- **100 influencers were paid** (likely with Russian money)

# WHAT WE ALSO KNOW

 CORRECTIV

## FACT-CHECKING

### Inside Doppelganger – How Russia uses EU companies for its propaganda

For two years, Russia has been spreading its propaganda using cloned websites of Western media outlets. The campaign known as "Doppelganger" hasn't been stopped yet – despite being sanctioned by the EU and US. A CORRECTIV investigation found that European companies are part of the internet infrastructure that keeps the campaign running.

von [Max Bernhard](#), [Alexej Hock](#), [Sarah Thust](#)  
22. July 2024



The Russian disinformation campaign Doppelganger has a global reach. CORRECTIV uncovered the internet infrastructure required to keep it running. (collage: Ivo Mayr / CORRECTIV)

- The **methods** (especially of right-wing parties) have been known for a long time
- **Coordinated influence campaigns** are standard tools of various foreign intelligence services
- The influence of social media on voting behavior is often **overestimated**

# TAKEAWAYS

- the **digital services act** marks a new area in platform accountability
- the impact of recommender systems in socio-technical systems is hard to assess
- **interdisciplinary** research is necessary to study social media platforms

Follow our work: [ai-forensics.org](https://ai-forensics.org), [tiktok-audit.com](https://tiktok-audit.com)

