

An investigative method to  
measure content on TikTok





**Carmen Aguilar García**

Data journalist  
@c\_aguilargarcia



**Zeke Hunter-Green**

Software developer  
@zekehg

# The Data Projects team



**Pamela Duncan**  
Data Projects Editor



**Carmen Aguilar García**  
Data Journalist



**Michael Goodier**  
Data Journalist

# Digital Investigations team



**Luke Hoyland**  
Product Manager



**Sabina  
Bejasa-Dimmock**  
Engineering Manager



**Zeke Hunter-Green**  
Engineer



**Daniel Hugenroth**  
Engineer



**Sam Cutler**  
Engineer



**Dominic Kendrick**  
Engineer

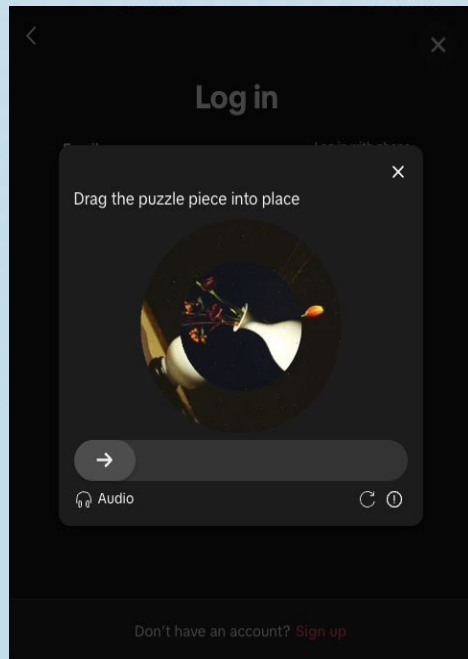
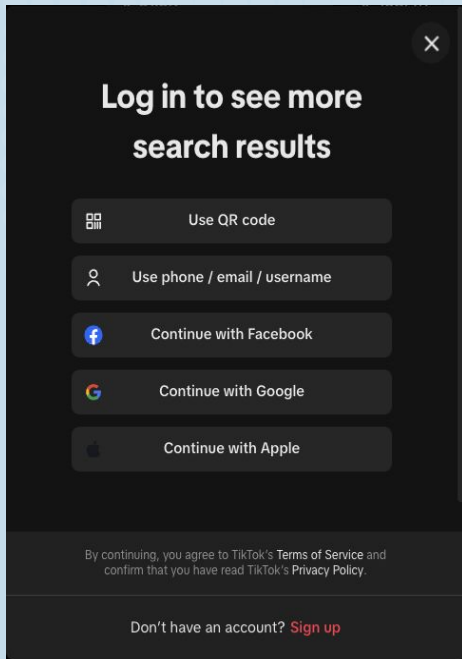


**Phil McMahon**  
Engineer

**Warning!**  
**Not a bed of roses**

# Researching TikTok is very difficult!

1. Opaque algorithm
2. Legal issues
3. Bot blocking



**So then... Why TikTok?**



# Very influential

(especially for young people)

< 2024

11/27/24 • COMMUNICATION

## One in five young people spends a dangerous amount of time on TikTok, more than two hours a day

*A study suggests carrying out regular audits of TikTok's algorithm in order to prevent its addictive effects*

[Nature](#)

“Having done comparative analyses, we’re really struck by just how **front-and-center youth identities** are on TikTok,” [Dr Literat in the NYT](#)

7 November 2023

**Global: TikTok’s ‘For You’ feed risks pushing children and young people towards harmful mental health content**

[Amnesty International](#)



# News consumption

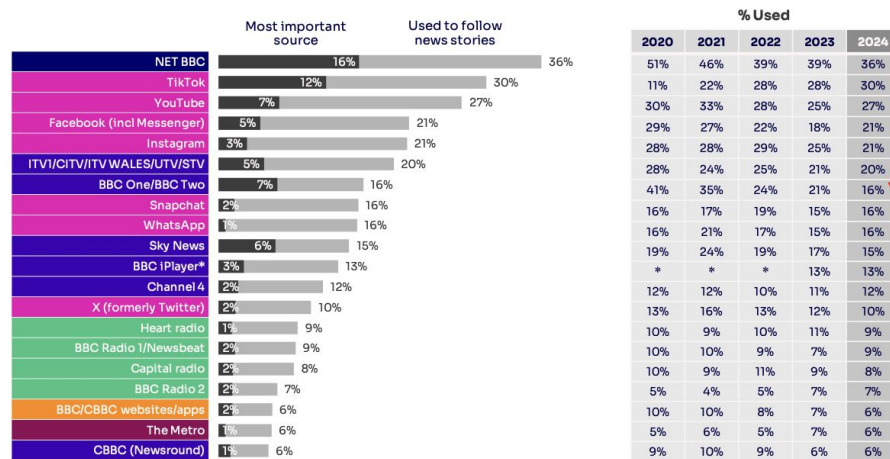
Figure 4: Top 20 news sources: 2019-2024<sup>5</sup>

	2019	2020	2022	2023	2024
BBC One	58%	56%	53%	49%	43% ▼
ITV1/ITV WALES/UTV/STV	40%	41%	35%	34%	30% ▼
Facebook	35%	34%	32%	30%	30%
BBC iPlayer	-	-	-	24%	23%
YouTube	6%	6%	8%	7%	19% ▲
Sky News Channel	23%	25%	23%	21%	19%
BBC News Channel	23%	21%	24%	23%	18% ▼
BBC news online (website/app)	25%	23%	23%	22%	18% ▼
Instagram	13%	14%	16%	16%	18%
X (formerly known as Twitter)	16%	17%	17%	17%	15%
WhatsApp	14%	13%	14%	13%	14%
Channel 4	17%	18%	17%	16%	14%
Google (search engine)	19%	15%	12%	13%	14%
Daily Mail/Mail on Sunday	18%	17%	15%	14%	13%
Channel 4 / Channel 4+ (All 4/All4+)	-	-	-	8%	12% ▲
ITVX/ITVX Premium	-	-	-	11%	12%
TikTok	-	1%	7%	10%	11%
The Guardian/Observer	11%	10%	10%	10%	10%
BBC Radio 2	12%	12%	11%	11%	10%
BBC Sounds	-	-	-	-	9%

▲ ▼ Significantly higher/lower than 2023

Ofcom news consumption survey 2024. Question: D2a-D8a. Thinking specifically about <platform>, which of the following do you use for news nowadays? D8a Question wording updated for news brands online. Base: All adults 16+ - 2024=5466, 2023=4556, 2022 W2=2792, 2020=4576, 2019=4691.

Figure 18: Top news sources used by 12-15s: 2020-2024



▲ ▼ Significantly higher/lower than 2023

Source: Ofcom Teens News Consumption Survey 2024 Question: D1. Which, if any, of the following do you use to read, watch, listen to or follow news stories? E1a. Here is a list of all the news sources that you say you use. Which ONE of these is the most important news source to you? Base: All teens aged 12-15 – 2024=1007, 2020-2023=1000-1010 \*BBC iPlayer added in 2023.

# Political tool

## TikTok Is Shaping Politics. But How?

Two researchers have studied political expression on the app since the Musical.ly era. Here's what they found.

[New York Times](#)

## How TikTok is shaping US democracy – swipe by swipe

“TikTok in 2022 is what Facebook was for the Obama era – if candidates want to **connect with young voters** (...), they can't afford not to be on it.”

[Heinrich Boll Stiftung](#)



## *The TikTok Star and Political Chameleon Vying for Colombia's Presidency*

Rodolfo Hernández bills himself as a paragon of democracy and a successful businessman who cares for the poor. A trip to Bucaramanga, the mountain-fringed city where he built his empire, reveals a different picture.

[New York Times](#)

# Political tool

NEWS > POLITICS

## Europe's far right uses TikTok to win youth vote

Gen Z influencers inspire right-wing leaders like France's Jordan Bardella ahead of June's EU election.

[Politico](#)



[FT](#)

POLITICS | GERMANY

## AfD eyes young voters in Germany for EU election boost

Stephanie Höppner  
03/03/2024

German 16- and 17-year-olds will be allowed to vote in the upcoming European Parliament elections. As a result, campaigners are increasingly turning to video-sharing platform TikTok.

[Deutsche Welle](#)

# Political tool

## Why the AfD dominates TikTok – and what could help against it

**“Their (the AfD) political messages are at least simple and emotional, and often massively shortened and misleading. In other words: They fit perfectly with the functional logic of social media”**

[Social Media Watchblog](#)

How far-right parties seduced young voters across Europe

Albena Azmanova



Young people abandoned mainstream parties in last weekend's European elections - a sign that their anxieties are not being heard

[The Guardian](#)

# Political tool



# **The Guardian's methodology**



# The Guardian's methodology

1. Gathering content from TikTokers
2. Searching and gathering content
3. Researching “For You” feed



# Gathering content from TikTokers

General election 2024

This article is more than 10 months old

## Nigel Farage outperforms all other UK parties and candidates on TikTok

Exclusive: Videos on Reform leader's account show more engagement and average views than any other candidate

UK election live - latest updates

Carmen Aguilar García,  
Michael Goodier, and  
Pamela Duncan

Wed 26 Jun 2024 06:00 BST



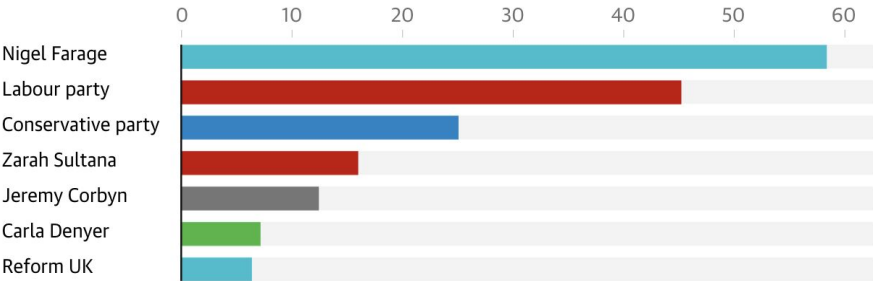
Share



Nigel Farage in Newton Abbot on Monday. His Reform UK voters via social media and podcast appearances. Photograph

### Farage and Labour have had the most engagement on TikTok

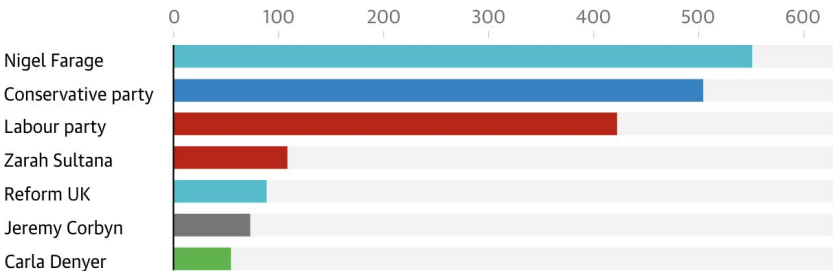
Number of likes, shares or comments per video, thousands



Guardian graphic. Source: Guardian analysis of TikTok content published 22 May to 17 June inclusively

### Videos of Farage, the Tories and Labour have had the most views

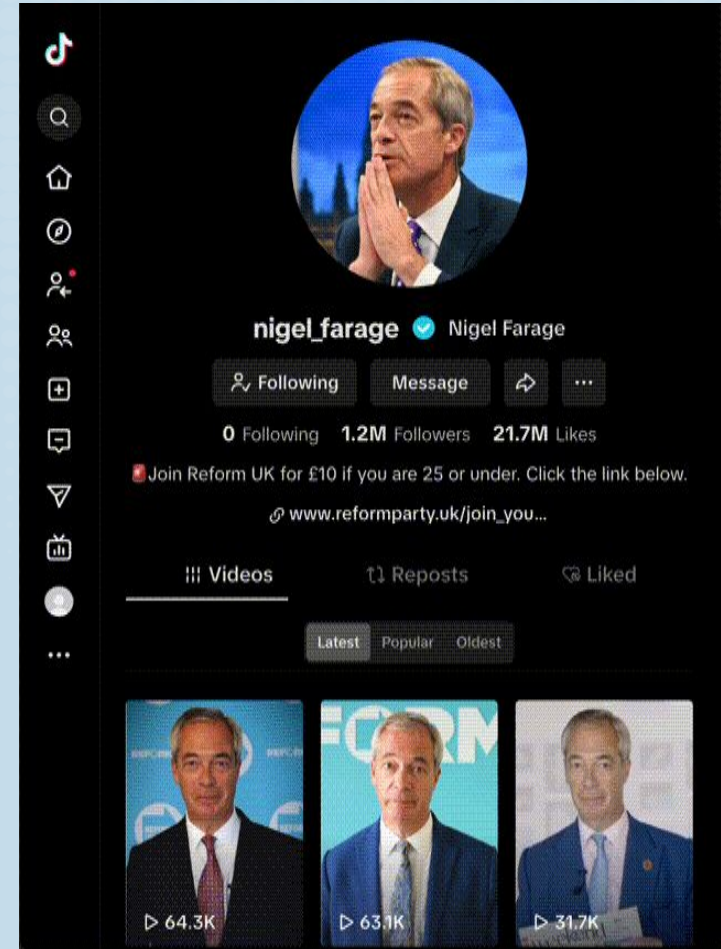
Number of view A bar graph showing the number of views per public video



Guardian graphic. Source: Guardian analysis of TikTok content published 22 May to 17 June inclusively

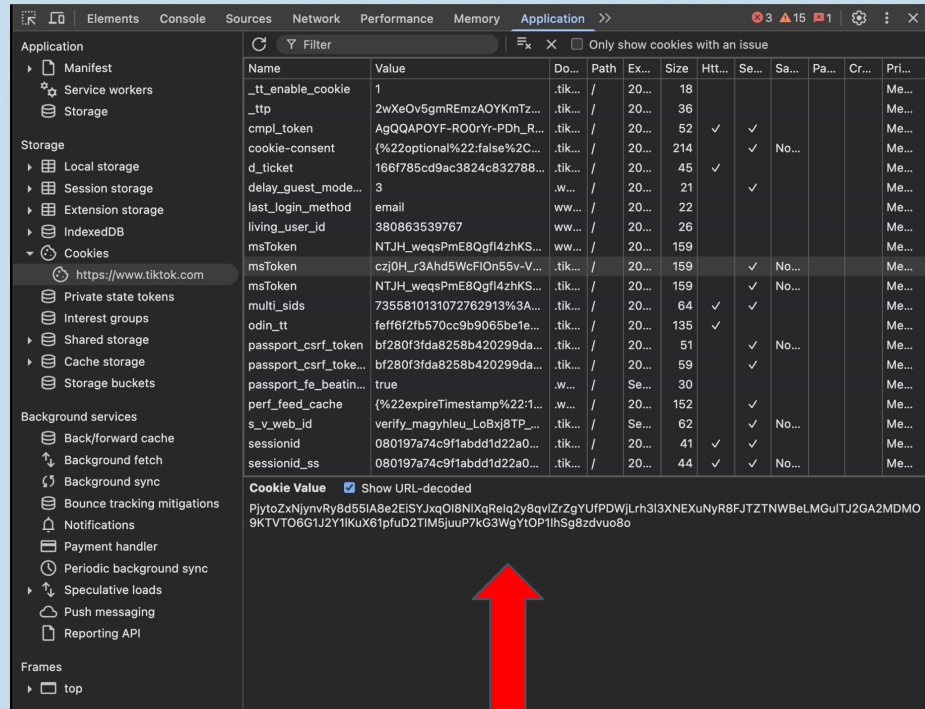
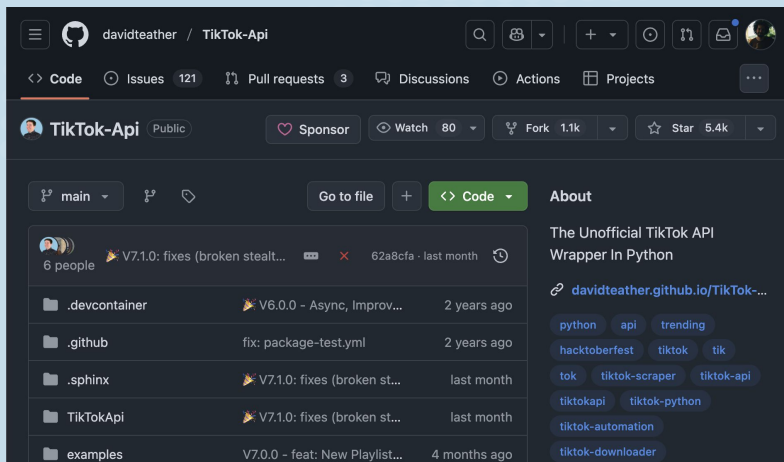
## Gathering content from TikTokers

- Manually search for TikTok accounts (3,200 candidates, 71% of all standing for election)
- Of that total, 500 TikTok accounts identified
- Gather all video content from those accounts
- Discard content published before the general election was called for the analysis



# Gathering content from TikTokers

- <https://github.com/davidteather/TikTok-API>
- Uses [Playwright](#) + undocumented API



```

1  import json
2  import os
3  from TikTokApi import TikTokApi
4
5
6  ms_token = os.environ.get("MS_TOKEN", None)
7  OUTPUT_DIR = "output/profiles"
8  INPUT_CSV = "input/profile_usernames.csv"
9
10
11  async def get_user_posts():
12      ... async with TikTokApi() as api:
13          ... await api.create_sessions(
14              ... ms_tokens=[ms_token],
15              ... num_sessions=1,
16              ... sleep_after=3,
17              ... headless=False,
18              ... )
19
20      ... author_username = "nigel_farage"
21      ... try:
22          ... user = api.user(author_username)
23          ... user_data = await user.info()
24          ... print(user_data)
25          ... video_data = []
26          ... async for video in user.videos(count=50000):
27              ... video_data.append(video.as_dict)
28
29          ... print(f"found {len(video_data)} posts")
30
31          ... with open(f"{OUTPUT_DIR}/{author_username}.json", "w") as f:
32              ... f.write(json.dumps(video_data))
33          ... except Exception as e:
34              ... print("error scraping", author_username, e)
35
36
37  if __name__ == "__main__":
38      ... import asyncio
39
40      ... asyncio.run(get_user_posts())
41

```

```

~/code/journalism-scrapers/tiktok-scraper gll:(main) v s1 A Pair M 1 2 Dispatch Ctrl 0 0 1
pipenv run python src/app/nigel.py Ctrl

```



```

562 },
563 {
564   "AIGCDescription": "",
565   "CategoryType": 120,
566   "author": {
567     "avatarLarger": "https://p16-common-sign-useast2a.tiktokcdn-us.com/tos-useast2a-avt-0068-euttp/3c671e78458f21b9707f0659380164e0~tplv-tiktokx-cropcenter:1080:1080.jpeg?dr=96486refresh_token=855f74d56x-expire=1748005200&x-signature=Umk2F0MvjP2vAtetC5wInRiha8KclX3D&t=4d5b04746ps=137406106shp=a5d480786shcp=81f88b706ldc=useast8",
568     "avatarMedium": "https://p16-common-sign-useast2a.tiktokcdn-us.com/tos-useast2a-avt-0068-euttp/3c671e78458f21b9707f0659380164e0~tplv-tiktokx-cropcenter:720:720.jpeg?dr=96486refresh_token=4bcb72b76x-expire=1748005200&x-signature=Wfbk2BnPmQjibgv7y6Xl3bXoDKo4e3D&t=4d5b04746ps=137406106shp=a5d480786shcp=81f88b706ldc=useast8",
569     "avatarThumb": "https://p16-common-sign-useast2a.tiktokcdn-us.com/tos-useast2a-avt-0068-euttp/3c671e78458f21b9707f0659380164e0~tplv-tiktokx-cropcenter:100:100.jpeg?dr=96486refresh_token=45988e085x-expire=1748005200&x-signature=gdufHFTuihhSiaaDKirwCLdIEto430&t=4d5b04746ps=137406106shp=a5d480786shcp=81f88b706ldc=useast8",
570     "commentSetting": 0,
571     "downloadSetting": 0,
572     "duetSetting": 0,
573     "ftc": false,
574     "id": "7049709431760110590",
575     "isADVirtual": false,
576     "isEmbedBanned": false,
577     "nickname": "Nigel_Farage",
578     "openFavorite": false,
579     "privateAccount": false,
580     "relation": 0,
581     "secUid": "MS4wLjABAAAA161f-DxghTgmTTFYZA1j1cPhTdJNBS74uDHv1UowP_1JU_GJ-wsG-qyISoyKC",
582     "secret": false,
583     "signature": "\ud83d\udea8Join Reform UK for \u00a310 if you are 25 or under. Click the link below.",
584     "stitchSetting": 0,
585     "uniqueId": "nigel_farage",
586     "verified": true
587   },
588   "authorStats": {
589     "diggCount": 74,
590     "followerCount": 1200000,
591     "followingCount": 0,
592     "friendCount": 0,
593     "heart": 21700000,
594     "heartCount": 21700000,
595     "videoCount": 403
596   },
597   "authorStatsV2": {
598     "diggCount": "74",
599     "followerCount": "1200000",
600     "followingCount": "0",
601     "friendCount": "0",
602     "heart": "21700000",
603     "heartCount": "21700000",
604     "videoCount": "403"

```

## Searching and gathering content

## General election 2024

🕒 This article is more than 9 months old

## Explainer

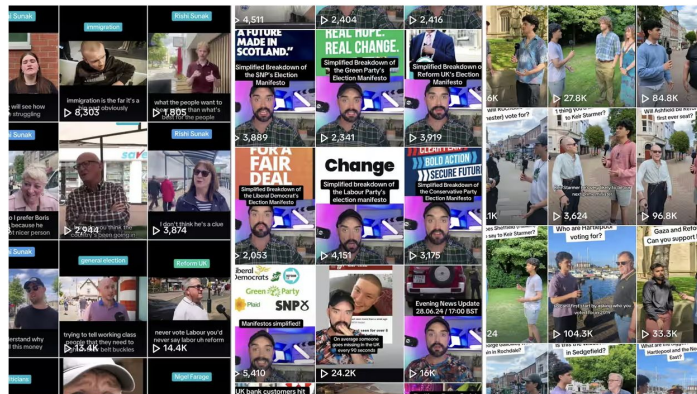
# #ukpolitics: how the 2024 general election has played out on TikTok

## Our sample of political content shows how campaigns and individuals have been using the video-sharing platform

- **UK general election live: latest updates**
- **Live tracker: UK general election results 2024**

**Carmen Aguilar García,  
Pamela Duncan, Michael  
Goodier and Zeke Hunter-  
Green with videos by Elena  
Morresi**

Thu 4 Jul 2024 14.08 BST



📹 TikTok creators are well aware that there is an audience for political content on the platform.  
Photograph: TikTok composite

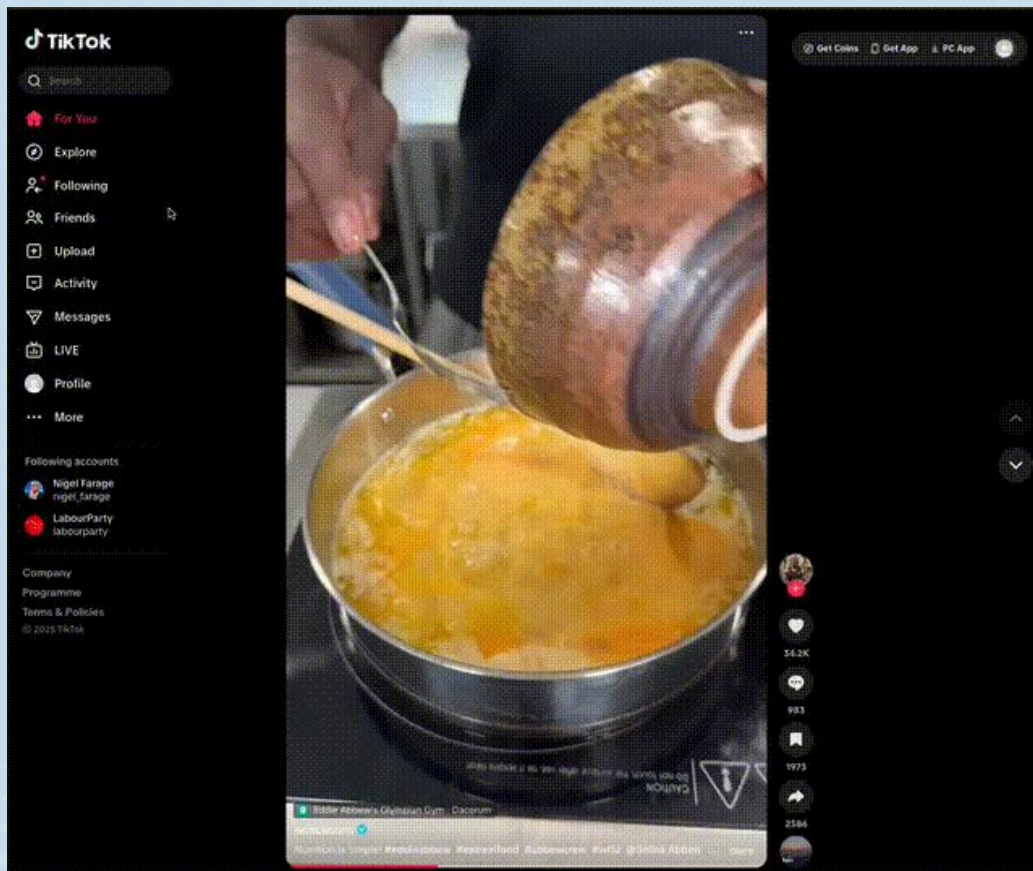
# The Guardian

## Searching and gathering content

- Select keywords and hashtags
- Create new accounts (even in different countries)
- Collect content several times a day at different times during ten days
- Clean the content (duplications, time frame, not relevant content creators)
- Refine sample of creators



## Hashtag search



## Schedule hashtag search

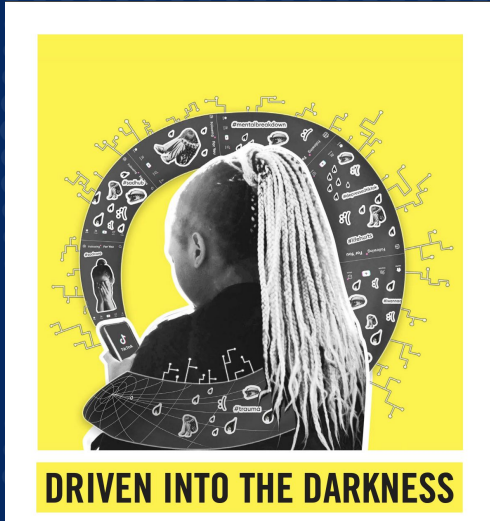
- Scheduling hashtag search
- Run script 4 times / day across the day  
in local time for ~2 weeks

```
117  monday_schedule = [6, 10, 14, 18]
118  tuesday_schedule = [7, 11, 15, 19]
119  wednesday_schedule = [8, 12, 16, 20]
120  thursday_schedule = [9, 13, 17, 21]
121  friday_schedule = [10, 14, 18, 22]
122  saturday_schedule = [11, 15, 19, 23]
123  sunday_schedule = [8, 12, 16, 20]
124
```

```
hashtag_search.py > ...
1  import TikTokApi
2
3
4  async def get_hashtag_videos(hashtag, ms_token):
5      ... async with TikTokApi() as api:
6          ... await api.create_sessions(
7              ...     ms_tokens=[ms_token],
8              ...     num_sessions=1,
9              ...     sleep_after=3,
10             ...     headless=True,
11             ... )
12             ... tag = api.hashtag(name=hashtag)
13             ... posts = []
14             ... async for v in tag.videos(count=500):
15                 ... posts.append(v.as_dict)
16             ... return posts
17
```

```
140
141  schedule = Scheduler()
142  for hashtag in hashtags:
143      ... for h in monday_schedule:
144          ... schedule.weekly(Monday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
145      ... for h in tuesday_schedule:
146          ... schedule.weekly(Tuesday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
147      ... for h in wednesday_schedule:
148          ... schedule.weekly(Wednesday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
149      ... for h in thursday_schedule:
150          ... schedule.weekly(Thursday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
151      ... for h in friday_schedule:
152          ... schedule.weekly(Friday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
153      ... for h in saturday_schedule:
154          ... schedule.weekly(Saturday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
155      ... for h in sunday_schedule:
156          ... schedule.weekly(Sunday(dt.time(hour=h), get_hashtag_videos(hashtag, token)))
157
```

## Researching the ForYou feed



“20 minutes into the experiment ... 72% of the videos recommended in the next 40 minutes related to mental health struggles, with at least five references to suicidal thinking or the content creator’s death wish”

Amnesty International

● This article is more than 2 years old

### How TikTok bombards young men with misogynistic videos

Observer investigation shows how online platform's algorithm pushed Andrew Tate posts to an imaginary teenager

● **Inside the violent, misogynistic world of TikTok's new star, Andrew Tate**

“When opening the app again a week later, the account was again flooded with Tate content, with eight out of the first 20 videos being of Tate.

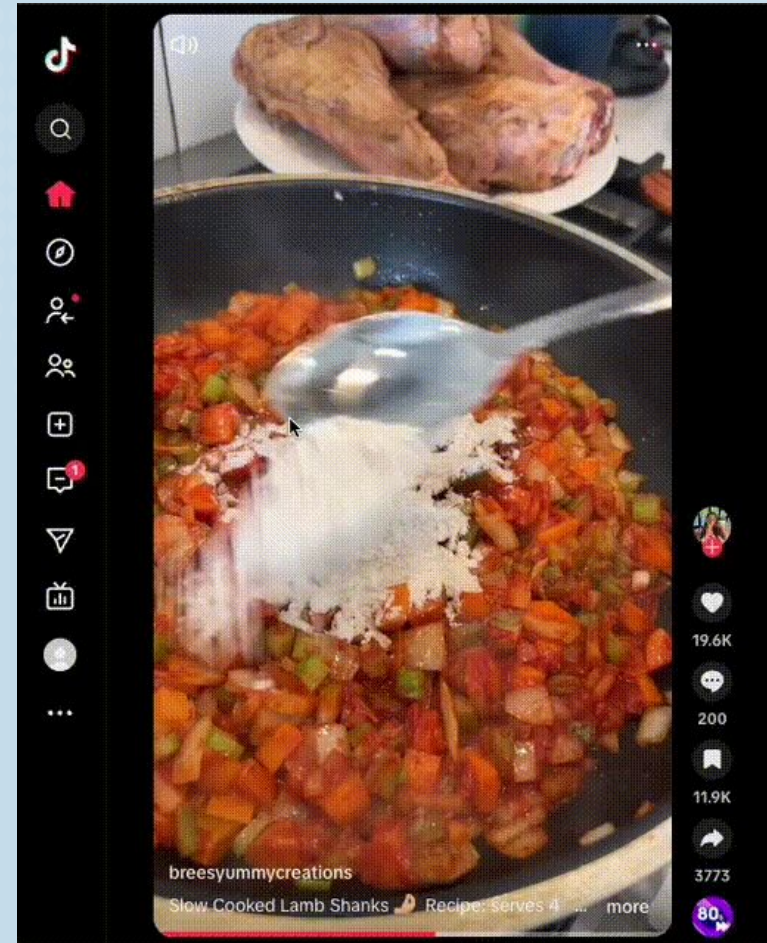
The algorithm also suggested videos from Dr Jordan Peterson, a Canadian psychologist known for his rightwing views; men’s coaching programmes and videos from men’s rights activists.”

The Observer



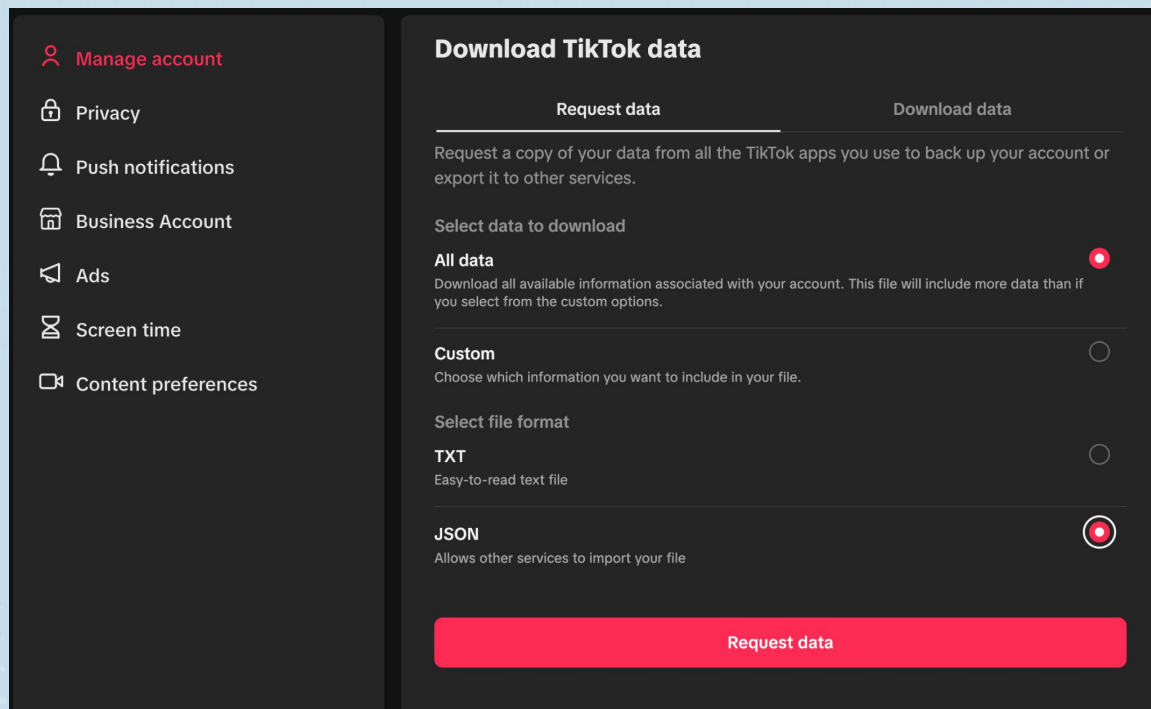
## Researching the ForYou feed

- Set up a new account
- Select relevant TikTokers and follow them
- Monitor TikTok for a few hours during several days, only watching content relevant to the investigation and skipping the rest
- Download your data



## Researching the “For You” feed

# Settings > Manage account > Download your data



## Researching the “For You” feed

“Download your data”

```
2      "Your Activity": {
88  ✓      "Watch History": {
89  ✓      "VideoList": [
90  ✓      {
91      "Date": "2024-03-01 16:24:41",
92      "Link": "https://www.tiktokv.com/share/video/7309502383174782251/"
93      },
94  ✓      {
95      "Date": "2024-03-01 16:25:29",
96      "Link": "https://www.tiktokv.com/share/video/731458362906001056/"
97      },
98  ✓      {
99      "Date": "2024-03-01 16:25:36",
100     "Link": "https://www.tiktokv.com/share/video/7317726700253482246/"
101     },
102  ✓      {
103     "Date": "2024-03-01 16:26:15",
104     "Link": "https://www.tiktokv.com/share/video/7315074280893582594/"
105     },
106  ✓      {
107     "Date": "2024-03-01 16:26:24",
108     "Link": "https://www.tiktokv.com/share/video/7340386780635565344/"
109     },
110  ✓      {
111     "Date": "2024-03-01 16:28:07",
112     "Link": "https://www.tiktokv.com/share/video/7325651045500128518/"
113     },
114  ✓      {
115     "Date": "2024-03-01 16:28:26",
116     "Link": "https://www.tiktokv.com/share/video/7262071434699099393/"
117     },
118  ✓      {
119     "Date": "2024-03-01 16:28:33",
120     "Link": "https://www.tiktokv.com/share/video/7340673010384342304/"
121     },
122  ✓      {
123     "Date": "2024-03-01 16:28:34",
```

## Setting up accounts in different countries and search hashtags



On one side of the border, users saw cat videos, on the other, tanks and plumes of smoke from the conflict. Image: Courtesy of NRK



## Hashtag search from different regions

- Creating accounts
  - New email address per user
  - Account creation using VPN
  - Some countries require registration in app

	A	B	C	E	G
1	country	ms_token	email address	tiktok handle	date of birth
2	UK	e7etiY-eVTm5U_Xsig3OvMbhnPjyHFoaqM5y0e	charleybishop65@gmail.com	charleyb82	March 10 1991
3	Malaysia	E7AO2NBYoWi8x_JQj6P01s2j4tdCNztzLCbDpp	charleybishop65+mys@gmail.com	n/a	Feb 4 1995
4	US	CkR1kjlWkXkxLIGKYjS_sFtmNYFZFfw4VJA3T5T	charleybishop65+usa@gmail.com	n/a	Jan 8 1990
5	Singapore	UzF9yY9_DivngmP3gzCTDmgzrV-5D6fLI4O7ysl	samb88960@gmail.com	n/a	July 10 1990
6	Egypt	YWMSeidKlVnHYisPcq5fF3REHyMPziXLTw_9F	amonisis04egypt@proton.me		Jun 5 1990
7	Saudi Arabia	FfZPCKYG69sAa9ErotZaaylHTRhgy9TS2Jy6G	aliibrahimsa93@outlook.com	abrahimsa30	
8	Turkey	rlt6BSOM6SFbCQzjPHKuzD0iv1vACOFaRWzIF	ali.muhammet1351@gmail.com	amuhammet52	
9	United Arab Emirates	OKBRuI9SCSsApp_DRWjLthLU_fejgQ16PeyNH	samira.ziad@protonmail.com	sziad84	Feb 10 1994
10	Indonesia	ZByvgYy_VjEkWYrlnwkNAIKEYJ-uEHfIRLNu3dg	riogarcia957@gmail.com	riogarcia957	6 October 1992
11	Australia	3wO5YMJiFqYSWxYdKTqA7JmcVOxCnFPkTsQ	1989samtaylor@gmail.com	1989samtaylor	11 September 19

## Hashtag search from different regions

- Decodo / smartproxy residential proxies



Authentication interface for Decodo / smartproxy residential proxies.

**AUTHENTICATION**: [Redacted] v

**PASSWORD**: [Redacted] [Eye icon] [Refresh icon]

**USED TRAFFIC**: 1.2 GB [Gear icon]

**LOCATION**: Random v

**ZIP CODE**: USA only [Flag icon]

**SESSION TYPE**: Sticky (10min) v

```
curl -U "my.decodo.com:30001" -x "https://ip.decodo.com/json"
{
  "browser": {
    "name": "",
    "version": ""
  },
  "platform": {
    "os": "undefined undefined"
  },
  "engine": {},
  "isp": {
    "isp": "TM Net",
    "asn": 4788,
    "domain": "",
    "organization": "TM Net"
  },
  "city": {
    "name": "Johor Bahru",
    "code": "01",
    "state": "Johor",
    "time_zone": "Asia/Kuala_Lumpur",
    "zip_code": "81300",
    "latitude": 1.5438,
    "longitude": 103.6461
  },
  "proxy": {
    "ip": "115.134.106.82"
  },
  "country": {
    "code": "MY",
    "name": "Malaysia",
    "continent": "Asia"
  }
}
```

## Researching the “For You” feed

- Answers questions like “Do Tiktok accounts based in Ukraine see more content related to the war than those based in Russia?”
- Need to decide on
  - independent variables (location of TikTok user)
  - dependent variables (# war-related posts in “For you” feed)
  - control variables (gender, age, interests, etc.)

## Data and its limitations

- Only public videos
- No deleted or deactivated accounts
- We don't know the algorithm
- Analysis based on frequency, not sentiment
- Metrics used: view rate and engagement rate

**In summary...**

## In summary...

- It is difficult and it changes all the time
- It is experimental
- Find good experts – our recommendations:



<https://ati.io/>

A<sub>I</sub> FORENSICS

<https://aiforensics.org/>

- We want to hear from others

Thanks

