53,473 views

911 0 **y** 

+ My Authors



THREAD: THE TWITTER FILES PART TWO.

Save as PDF

□ Bookmark

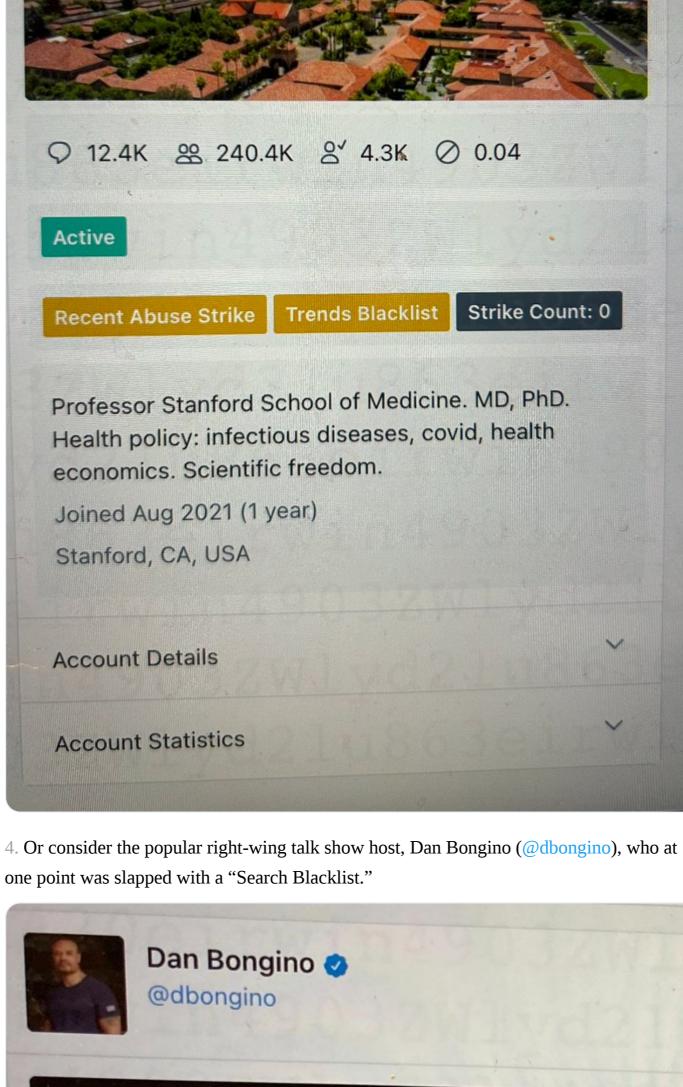
## TWITTER'S SECRET BLACKLISTS.

## 1. A new #TwitterFiles investigation reveals that teams of Twitter employees build blacklists, prevent disfavored tweets from trending, and actively limit the visibility of entire

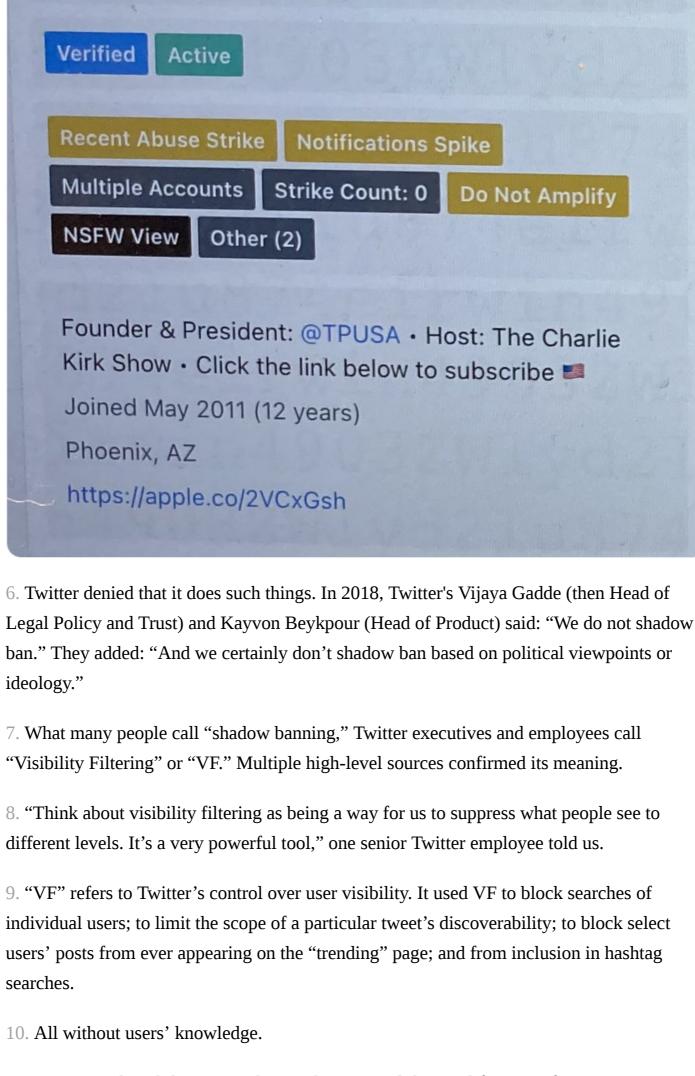
accounts or even trending topics—all in secret, without informing users. 2. Twitter once had a mission "to give everyone the power to create and share ideas and information instantly, without barriers." Along the way, barriers nevertheless were erected.

3. Take, for example, Stanford's Dr. Jay Bhattacharya (@DrJBhattacharya) who argued that Covid lockdowns would harm children. Twitter secretly placed him on a "Trends Blacklist," which prevented his tweets from trending.

Jay Bhattacharya @DrJBhattacharya







Response Team - Global Escalation Team, or SRT-GET. It often handled up to 200 "cases" a day. 13. But there existed a level beyond official ticketing, beyond the rank-and-file moderators following the company's policy on paper. That is the "Site Integrity Policy, Policy Escalation Support," known as "SIP-PES."

14. This secret group included Head of Legal, Policy, and Trust (Vijaya Gadde), the Global

Head of Trust & Safety (Yoel Roth), subsequent CEOs Jack Dorsey and Parag Agrawal, and

15. This is where the biggest, most politically sensitive decisions got made. "Think high

follower account, controversial," another Twitter employee told us. For these "there would

16. One of the accounts that rose to this level of scrutiny was @libsoftiktok—an account

that was on the "Trends Blacklist" and was designated as "Do Not Take Action on User

others.

be no ticket or anything."

Without Consulting With SIP-PES."

ATEGORY Testing

Communities

Guano

Email

Payments

policy against "hateful conduct."

Hateful Conduct policy." See here:

**Yoel Roth**09:02:09

need some attention.

**Yoel Roth**11:51:36

filtering."

@TheFP.

Matt Taibbi

Six TWITTER, THE FBI

**SUBSIDIARY** 

media content.

Read 45 tweets

 $\bigcirc$ 

1. THREAD: The Twitter Files, Part

2. The #TwitterFiles are revealing more

collects, analyzes, and flags your social

every day about how the government

3. Twitter's contact with the FBI was

Try unrolling a thread yourself!

Replying to @wrathofgnon

Replying to @wrathofgnon

**Bari Weiss** 

Oct 17, 2021

It's been quite a weekend... Last night I

@LAPressClub and @yudapearl. This

I was thrilled, also, to finally sit down

was deeply honored to receive the

forward.com/news/israel/47...

@bariweiss

Daniel Pearl Award from

is the speech I gave:

@threadreaderapp unroll

Wrath Of Gnon @wrathofgnon · Jan 6

17

 $\bigcirc$ 

In in Hualien County a private 6ha butterfly reserve is being used to also grow indigo plants, which were a major cash crop until about a century ago. Underneath the trees indigo plants provide food and shelter for the

butterflies: excess indigo leaves are harvested and sold.

12. The group that decided whether to limit the reach of certain users was the Strategic

① DO NOT TAKE ACTION ON USER WITHOUT CONSULTING WITH SIP-PES@ Timeline Libs of TikTok Spaces Identity @libsoftiktok Live Videos ! Email is protect accessed for I

Email is a private do

Last Tweet

2022-12-07 01:54

**Site Policy Recommendation** Site Policy recommends placing @LibsOfTiktok ([LTT] 1.3M followers, not verified) in a 7-day timeout at the account level [meaning, not for a specific Tweet] based on the account's of any of Site Policy's policies, we will move forward with permanent suspension. **Assessment** Since its most recent timeout, while LTT has not directly engaged in behavior violative of the Hateful Conduct policy, the user has continued targeting individuals/allies/supporters of the LGBTQIA+ community for alleged misconduct. The targeting of at least one of these institutions 20. The committee justified her suspensions internally by claiming her posts encouraged online harassment of "hospitals and medical providers" by insinuating "that genderaffirming healthcare is equivalent to child abuse or grooming." 21. Compare this to what happened when Raichik herself was doxxed on November 21, 2022. A photo of her home with her address was posted in a tweet that has garnered more than 10,000 likes. 22. When Raichik told Twitter that her address had been disseminated she says Twitter Support responded with this message: "We reviewed the reported content, and didn't find it to be in violation of the Twitter rules." No action was taken. The doxxing tweet is still up.

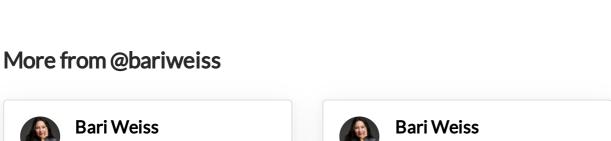
prevalence overall). We got Jack on board with implementing this for civic integrity in the near term, but we're going to need to make a more robust case to get this into our repertoire of policy remediations - especially for other policy domains. So I'd love research's POV on that. 25. Roth wrote: "The hypothesis underlying much of what we've implemented is that if exposure to, e.g., misinformation directly causes harm, we should use remediations that reduce exposure, and limiting the spread/virality of content is a good way to do that."

One of the biggest areas I'd \*love\* research support on is re: non-removal policy interventions like disabling engagements and deamplification/visibility filtering. The hypothesis underlying much of what we've implemented is that if exposure to, e.g.,

misinformation directly causes harm, we should use remediations that reduce exposure,

and limiting the spread/virality of content is a good way to do that (by just reducing

30. Watch @mtaibbi for the next installment. Missing some Tweet in this thread? You can try to force a refresh Tweet **f** Share **☑** Email **★** Keep Current with Bari Weiss Stay in touch and get notified when new Twitter may remove this content at anytime! unrolls are available from this author! Save it as PDF for later use! 2 Read all threads + Add to "My Authors" ■ Save this thread as PDF



## with my friend @benshapiro for his scho... Sunday Special. Watch our Read 4 tweets Read 7 tweets Did Thread Reader help you today? Support us! We are indie developers!

Become a Premium Member (\$3/month or \$30/year) and get exclusive features! ♥ Become Premium Don't want to be a Premium member but still want to support us? Make a small donation by buying us coffee (\$5) or help with server cost (\$10) P Donate via Paypal

Read 4 tweets

**Bari Weiss** @bariweiss

I'm grateful to everyone who is

amplifying his voice. Thank you

shows.acast.com/wethefifth/epi...

podcasts.podinstall.com/devil-may-

Thank you, @RubinReport:

Thank you, @megynkelly:

@wethefifth:

May 3, 2021

Paul Rossi is morally courageous and

THE DAN -

Q 53K № 1.9M & 192.8K Ø 0.17

11. "We control visibility quite a bit. And we control the amplification of your content quite a bit. And normal people do not know how much we do," one Twitter engineer told us. Two additional Twitter employees confirmed.

Q 8.8K 

2 1.6M 2 0.8K 

○ 0.09 Account is active Phone Lists High Profile Activity Direct Messages Last Access Twitter Blue Verified 2022-12-07 02:03 nteractions (!) IP addresses 2 Account purposes. ompromise All videos belong to their respective owners. Access IP addres submissions@libsoftiktok.com. DM submissions. (i) Spam Joined Nov 2020 (2 years)

Depths of Hell

Raichik was blocked from posting for as long as a week.

https://allmylinks.com/libsoftiktok

17. The account—which Chaya Raichik began in November 2020 and now boasts over 1.4

million followers—was subjected to six suspensions in 2022 alone, Raichik says. Each time,

18. Twitter repeatedly informed Raichik that she had been suspended for violating Twitter's

19. But in an internal SIP-PES memo from October 2022, after her seventh suspension, the

committee acknowledged that "LTT has not directly engaged in behavior violative of the

continued pattern of indirectly violating Twitter's Hateful Conduct Policy by tweeting content that either leads to or intends to incite harassment against individuals and institutions that support LGBTQ communities. At this time, Site Policy has not found explicitly violative Tweets, which would result in a permanent suspension of the account. This type of enforcement action [repeated 7-day timeouts at the account-level] will not lead to permanent suspension, however: should LTT engage in any other direct Tweet-level violations

Thanks for reaching out. We reviewed the reported content, and didn't find it to be in violation of the Twitter rules. In this case, no action will be taken at this If you have further concerns about intellectual property, your privacy, or your personal safety, the following guidelines can be of assistance: 23. In internal Slack messages, Twitter employees spoke of using technicalities to restrict the visibility of tweets and subjects. Here's Yoel Roth, Twitter's then Global Head of Trust & Safety, in a direct message to a colleague in early 2021:

A lot of times, SI has used technicality spam enforcements as a way to solve a problem created by Safety under-enforcing their policies. Which, again, isn't a problem per se - but it keeps us from addressing the root cause of the issue, which is that our Safety policies

24. Six days later, in a direct message with an employee on the Health, Misinformation,

Privacy, and Identity research team, Roth requested more research to support expanding

"non-removal policy interventions like disabling engagements and deamplification/visibility

26. He added: "We got Jack on board with implementing this for civic integrity in the near term, but we're going to need to make a more robust case to get this into our repertoire of policy remediations – especially for other policy domains." 27. There is more to come on this story, which was reported by @AbigailShrier

@ShellenbergerMD @NellieBowles @IsaacGrafstein and the team The Free Press

Keep up with this unfolding story here and at our brand new website: thefp.com.

28. The authors have broad and expanding access to Twitter's files. The only condition we

29. We're just getting started on our reporting. Documents cannot tell the whole story here.

A big thank you to everyone who has spoken to us so far. If you are a current or former

Twitter employee, we'd love to hear from you. Please write to: tips@thefp.com

A new media company built on the ideals that were once the bedrock of American journalism.

The Free Press

http://thefp.com

agreed to was that the material would first be published on Twitter.

People who liked this thread also liked...

Matt Taibbi

2.In July of 2020, San Francisco FBI

agent Elvis Chan tells Twitter executive

Yoel Roth to expect written questions

from the Foreign Influence Task Force

(FITF), the inter-agency group that

deals with cyber threats.

Read 14 tweets

1

1.THREAD: Twitter Files

Supplemental

Name Redacted @NameRedacted247

Dec 19

1. After learning that Twitter employs at

least 15 former FBI agents, I searched

Facebook. What I found is alarming

Facebook currently employs at least

115 people, in high-ranking positions,

FBI/CIA/NSA/DHS: 17 CIA 37 FBI 23

that formerly worked at

NSA 38 DHS Thread

Read 60 tweets

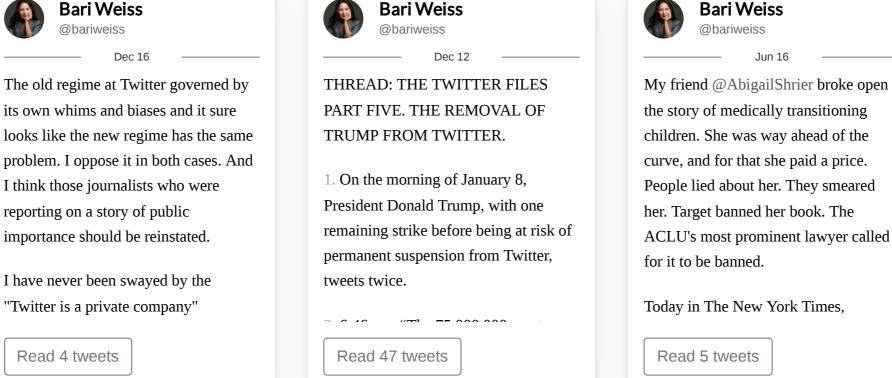
Follow @ThreadReaderApp to mention us!

Practice here first or read more on our help page!

@threadreaderapp unroll

2. From a Twitter thread mention us with a keyword "unroll"

2. All, but a few, of the former



**Bari Weiss** 

Sep 10, 2021

living publishing my columns. I started

it because I want to run the best opinion

reporting, and fresh insight. I think we

I didn't start a newsletter to make a

page in the country, full of scoops,

delivered on that goal for back-to-

bariweiss.substack.com/p/back-to-

school week (Thread!):

@bariweiss

This site is made by just two indie developers on a laptop doing marketing, support and development! Read more about the story.

Or Donate anonymously using crypto!

0xfe58350B80634f60Fa6Dc149a72b4DFbc17D341E

3ATGMxNzCUFzxpMCHL5sWSt4DVtS8UqXpi

**Ethereum** 

**Bitcoin** 

💙 💙 Thank you for your support! 💙 💜

Follow Us on Twitter!

Help | About | TOS | Privacy

**Tweet** 

**f** Share