

Profile header for Bari Weiss (@bariweiss) with 354K followers and 9 tweets.

THREAD: THE TWITTER FILES PART TWO.

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.

Twitter profile for Jay Bhattacharya (@DrJBhattacharya) showing account details, bio, and suspension status: 'Trends Blacklist'.

Twitter profile for Dan Bongino (@dbongino) showing account details, bio, and suspension status: 'Search Blacklist'.

Twitter profile for Charlie Kirk (@charliekirk11) showing account details, bio, and suspension status: 'Do Not Amplify'.

6. Twitter denied that it does such things. In 2018, Twitter's Vijaya Gadde (then Head of Legal Policy and Trust) and Kayvon Beykpour (Head of Product) said: "We do not shadow ban."

7. What many people call "shadow banning," Twitter executives and employees call "Visibility Filtering" or "VF."

8. "Think about visibility filtering as being a way for us to suppress what people see to different levels. It's a very powerful tool," one senior Twitter employee told us.

9. "VF" refers to Twitter's control over user visibility. It used VF to block searches of individual users; to limit the scope of a particular tweet's discoverability; to block select users' posts from ever appearing on the "trending" page; and from inclusion in hashtag searches.

10. All without users' knowledge.

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.

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

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 be no ticket or anything."

16. One of the accounts that rose to this level of scrutiny was @libsoffitok—an account that was on the "Trends Blacklist" and was designated as "Do Not Take Action on User Without Consulting With SIP-PES."

Screenshot of a Twitter profile for Libs of TikTok (@libsoffitok) with a 'DO NOT TAKE ACTION ON USER WITHOUT CONSULTING WITH SIP-PES' warning.

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.

18. Twitter repeatedly informed Raichik that she had been suspended for violating Twitter's policy against "hateful conduct."

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 Hateful Conduct policy." See here:

Site Policy Recommendation document snippet regarding enforcement actions for @libsoffitok.

20. The committee justified her suspensions internally by insinuating "that gender-affirming 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."

Screenshot of a Twitter support message: "Thanks for reaching out. We reviewed the reported content, and didn't find it to be in violation of the Twitter rules."

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:

Direct message screenshot from Yoel Roth: "A lot of times, I have used technicality spam enforcements as a way to solve a problem created by Safety under-enforcing their 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 filtering."

Direct message screenshot from Yoel Roth: "One of the biggest areas I'd love research support on is re: non-removal policy interventions like disabling engagements and deamplification/visibility filtering."

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."

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 @IsaacGrafein and the team The Free Press @TheFP.

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

The Free Press logo and website information: "A new media company built on the ideals that were once the bedrock of American journalism."

28. The authors have broad and expanding access to Twitter's files. The only condition we agreed to was that the material would first be published on Twitter.

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

30. Watch @mtaibbi for the next installment.

People who liked this thread also liked...

Grid of three tweet recommendations from Matt Taibbi and Name Redacted.

Try unrolling a thread yourself!

Instructions on how to unroll a thread using the @threadreaderapp app.

More from @bariweiss

Grid of tweets from Bari Weiss, including one about the old regime at Twitter and another about a newsletter.

Did Thread Reader help you today? Support us! We are indie developers!

Options to support the developers: Become a Premium Member, donate via PayPal, or donate via Ethereum/Bitcoin.