This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers visit https://www.djreprints.com.

https://www.wsj.com/articles/facebook-suppresses-political-movements-patriot-party-11634937358

Facebook Increasingly Suppresses Political Movements It Deems Dangerous

Reluctant to stifle viral content it helps amplify, the social-media giant is playing whack-a-mole with groups it believes are potentially harmful

By Jeff Horwitz Follow and *Justin Scheck* Follow
Oct. 22, 2021 5:19 pm ET

After the Jan. 6 Capitol riot, far-right activists launched an online campaign to form what they called a Patriot Party as an alternative to the Republican Party.

Facebook Inc. worked to kill it, citing information it said showed the movement was being pushed by white nationalists and self-styled militias who had worked to instigate the riot, according to internal company documents reviewed by The Wall Street Journal.

Facebook engineers made it harder for organizers to share Patriot Party content, restricted the visibility of groups connected to the movement and limited "super-inviters" from recruiting new adherents, according to a March review.

"We were able to nip terms like Patriot Party in the bud before mass adoption," said another memo.

The surgical strike was part of a strategy Facebook adopted early this year to stop what it calls "harmful communities" from gaining traction on its platform before they spread too far. Rather than just taking action against posts that violate its rules, or that originate with actors such as Russia-based trolls, Facebook began putting its thumb on the scale against communities it deemed to be a problem. In April, based on the same policy, it took aim at a German conspiracy movement called Querdenken.

Internal Facebook documents, part of an array of company communications reviewed by the Journal for its Facebook Files series, show that people inside the company have long discussed a different, more systematic approach to restrict features that disproportionately amplify incendiary and divisive posts. The company rejected those efforts because they would impede the platform's usage and growth.

The reality is that Facebook is making decisions on an ad hoc basis, in essence playing whack-a-mole with movements it deems dangerous. By taking on the role of refereeing public discourse, Facebook has strayed from the public commitment to neutrality long espoused by Chief Executive Mark Zuckerberg.

And because of the enormous size of its global user base—the latest count is about 2.9 billion—its decisions about whom to silence, with no public disclosure or right of appeal, can have great impact.

A protester holds a Patriot Party flag outside the Washington state capitol on Feb. 6.

PHOTO: JOHN RUDOFF/ANADOLU AGENCY/GETTY IMAGES

The issue sits in the middle of one of the most sensitive debates around Facebook. Activists on the left have been urging the company to more forcefully block what they see as harmful content. Other activists, largely on the right, accuse Facebook of censorship directed at conservative voices.

Matt Perault, a former director of global public policy for the company who left for an academic post at Duke University in 2019, said documents shared with him by the Journal suggest that Facebook's commitment to being a neutral platform was slipping.

"It's understandable that the immense pressure on tech companies would push them to develop aggressive solutions to combat misinformation," Mr. Perault said. "But predictive, behavioral censorship seems fraught. In the absence of data suggesting otherwise, I think it's appropriate to be skeptical that the benefits will outweigh the costs."

Facebook spokesman Drew Pusateri acknowledged the tension in the company's work to combat dangerous viral social movements. "To find those solutions, we've had to invent new technologies and balance difficult trade-offs that society has struggled with for a long time, and without needed guidance from lawmakers and regulators," he said. "We know our solutions will never be perfect, but stories like these exist precisely because we confronted our toughest problems head-on."

A senior security official at Facebook said the company would seek to disrupt on-platform movements only if there was compelling evidence that they were the product of tightly knit circles of users connected to real-world violence or other harm and committed to violating Facebook's rules.

"When you start to think about authentic people organizing on the platform and contributing to harm, you have to be very careful," the official said in an interview. "It's very easy to get into a place where you're picking and choosing. And that isn't fair."

One challenge for the company has been balancing concern about fairness with recent history, in which groups such as foreign trolls and small conspiracy movements have used Facebook to get their message out to millions of people.

The Patriot Party was a loose and fast-growing collection of people and groups that had supported Donald Trump's false claim that last year's presidential election was stolen from him. In the aftermath of the Jan. 6 riot at the U.S. Capitol, Facebook and other social-media platforms worked to suppress many such groups, especially those associated with the "Stop the Steal" movement. In response, thousands of members rallied behind the Patriot Party name, starting Facebook groups, websites and chapters around the country.

Facebook employees watched the Patriot Party movement grow in real time. The company's automated systems showed that conversations about the proposed pro-Trump political movement were disproportionately heavy on hate speech and incitement to violence, and the company's researchers had spotted links between the party's promoters and armed movements, the documents show.

"We need to organize our militia to meet up with local police and weekend warriors. Wars are won with guns...and when they silence your commander in chief you are in a war," one member wrote on Jan. 9 on a Facebook Patriot Party page. It has since disappeared from Facebook, but was archived by an advocacy group, the Tech Transparency Project.

Still, nothing about the idea of founding a new political party itself broke Facebook's rules.

Larry Glenn, an Ohio man who is treasurer of the American Patriot Party of the U.S., said his group was growing early this year largely because of Facebook. The platform let it spread its message and directed like-minded people to it, some of whom were in turn directed to the party's website. "The website was growing quickly. We had a lot of followers," Mr. Glenn said in an interview.

Then Facebook started warning the group, according to Mr. Glenn, about divisive content posted to their group. He said his colleagues worked to take down rule-violating posts, but the group was still removed from Facebook. "They pulled the rug out from under us," he said. Now, the party is dormant, he added.

The targeted approach has been controversial within Facebook. Documents show employees have long championed "content agnostic" changes to the platform, which were both technically easier and less likely to raise free-speech concerns.

One such employee has been Kang-Xing Jin, who has been friends with Mr. Zuckerberg since their first day at Harvard University and now leads the company's health-related initiatives. In late 2019, Mr. Jin warned colleagues: If the company didn't dial back on automated recommendations and design features that disproportionately spread false and inflammatory posts, what he called "rampant harmful virality" could undermine Facebook's efforts to prevent the spread of toxic content before the 2020 election.

"There's a growing set of research showing that some viral channels are used for bad more than they are used for good," he wrote in one note. Facebook wouldn't have to eliminate virality to deal with the problem, he said, just dial it back.

Those suggestions were met with praise from other employees, the documents show, but generally didn't get traction with executives, leaving the ad hoc approach as the company's main weapon. A Facebook spokesman said the company took such concerns seriously, and that it adopted a proposal Mr. Jin championed to stop recommending users join groups related to health or political topics, and had taken steps to slow the growth of newly created groups.

"Provocative content has always spread easily among people," said Mr. Pusateri, the spokesman. "It's an issue that cuts across technology, media, politics and all aspects of society, and when it harms people, we strive to take steps to address it on our platform through our products and policies."

Facebook's trillion-dollar business is built largely on its unique ability to keep users coming back, in part by maximizing the viral spread of posts that people will share and reshare. Mr. Zuckerberg has often highlighted the benefits of virality. He points to the "ice-bucket challenge," in which thousands of users in 2014 filmed themselves pouring freezing water over their heads to raise money for charity.

Executives were slow to think about the downsides and what to do about them. "For the first 10 years of the company, everyone was just focused on the positive," Mr. Zuckerberg said in a 2018 interview with Vox.

The 2016 U.S. election changed that. Revelations of foreign interference, bot networks and false information left the company scrambling to identify how its platform could be abused and how to prevent it.

Its researchers found that company systems automatically and disproportionately spread harmful content, the internal documents show. Whatever content was shared, Facebook would recommend and spread a more incendiary mix.

Facebook took action against a German movement called Querdenken, whose supporters are shown protesting in November.

PHOTO: OMER MESSINGER/GETTY IMAGES

Particularly troublesome were heavy users, the kind of voices that Facebook's algorithm had long helped amplify. In at least nine experiments and analyses beginning in 2017, Facebook researchers found links popular with heavy users were disproportionately associated with false information and hyperpartisan content, the documents show.

"Pages that share misinformation tend to post at much higher rate than other pages with similar audience size," one research note states.

Researchers also found that efforts to boost the "relevance" of Facebook's News Feed were making the platform's problems with bad content worse.

In dozens of experiments and analyses reviewed by the Journal, Facebook researchers, data scientists and engineers found viral content favored conspiracy theories, hate speech and hoaxes. And they discovered that as the speed and length of the sharing chain grew, so did the odds that content was toxic.

To demonstrate to colleagues how Facebook's dynamics ended up promoting a toxic brew, a researcher created a Facebook account for a fictional 41-year-old named Carol Smith, according to a document describing the experiment. The researcher made Ms. Smith a "conservative mom" from Wilmington, N.C., interested in "young children, parenting, Christianity, Civics and Community." Her tastes leaned right but mainstream.

On the first day of the experiment, Facebook recommended humorous memes and generally conservative groups, which the fictional Ms. Smith joined. By the second day, it was recommending almost exclusively right-wing content, including some that leaned toward conspiracy theories. By the fifth day, the platform was steering Ms. Smith toward groups with overt QAnon affiliations. Selected content included false claims of "white genocide," a conspiracy-theory "watch party" and "a video promoting a banned hate group."

A subsequent study of the platform's recommendations to a liberal user found a similar distorting effect.

The documents show that employees pushed for the company to confront its reliance on virality, but that its leaders resisted.

One engineer in 2019 suggested killing the reshare button, which let users quickly spread posts across Facebook. Other suggestions were more incremental: to stop promoting reshared content unless it was from a close friend of the user; to moderately slow the platform's fastest-moving posts; or to lower the limit on daily group invitations from 2,250 a day.

Facebook data scientists intensified their scrutiny of viral problems during the 2020 election run-up, putting in place measures to analyze how fast harmful content was spreading. They found the company had inadvertently made changes that worsened viral problems.

In Facebook's internal communications system, called Workplace, Mr. Jin, Mr. Zuckerberg's former schoolmate, said research suggested Facebook was in the wrong part of the "virality tradeoff curve." He championed measures to damp virality on the platform. Many colleagues agreed that one attractive part of his suggestion was that slowing down the spread of viral information would affect everyone, no matter where they are on the political spectrum. It would help the company avoid the accusations of bias that comes when it targets a specific group.

Mr. Jin ran into skepticism from John Hegeman, Facebook's head of ads. At the time, Mr. Hegeman oversaw recommendations in Facebook's News Feed. He agreed with Mr. Jin's assertion that Facebook's systems appeared to magnify its content problems. But he contended that most viral content is OK, and asked whether it would be fair, or wise, to cut back.

"If we remove a small percentage of reshares from people's inventory," he wrote, "they decide to come back to Facebook less."

Facebook didn't follow Mr. Jin's advice. By early 2020, executives responsible for Facebook's election preparations were growing worried, the documents indicate. Facebook lacked even "a minimal level of reactive protection" against viral falsehoods when the fourth quarter began, according to one document, but the company wasn't prepared to change course.

The company moved from crisis to crisis in 2020. The platform boosted divisive material from QAnon conspiracy theorists, violent armed groups and the Stop the Steal movement, according to the internal company analyses.

In each case, the documents indicate, Facebook's tools turbocharged the growth of those movements, and the company stepped in to fight them only after they led to real-world violence or other harms. That thrust it repeatedly into messy arguments about whether its controls over speech on the platform were insufficient or overbearing and biased.

The company tried slowing its platform, but only as a temporary, emergency response, part of what Facebook referred to internally as "Break the Glass" measures. It first did so when false claims of election fraud took off in the immediate wake of the U.S. presidential election, then after the Jan. 6 riot.

In most cases, the company rolled those measures back afterward, according to the documents.

In the days following the Capitol riot, critics accused social-media companies of fueling the insurrection. At Facebook, systems showed that conversations about the Patriot Party movement were disproportionately heavy on hate speech and incitement to violence. Some employees, though, worried that taking aim at a grass-roots movement would mean that Facebook was tipping the political scales.

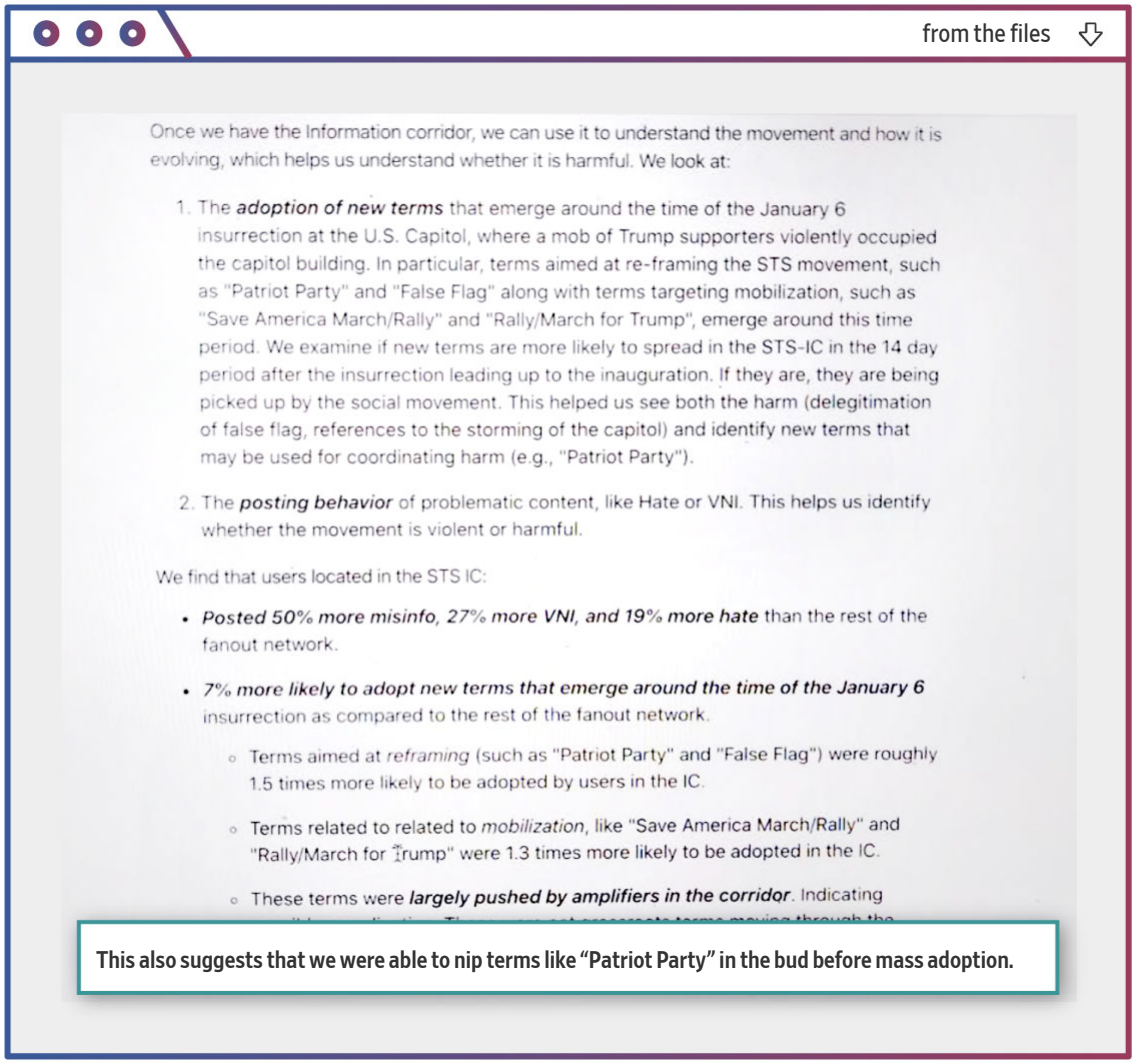
Some movements are following Facebook's rules but also spreading content the company deems "inherently harmful and violates the spirit of our policy," one Facebook researcher wrote. "What do we do when that authentic movement espouses hate or delegitimizes free elections? These are some of the questions we're trying to answer."

One Patriot Party supporter was Dick Schwetz, a Pennsylvania salesman and member of a chapter of the Proud Boys, a violent, far-right group, who said in an interview he was banned from Facebook before the presidential election. Mr. Schwetz later promoted the Patriot Party.

Dick Schwetz, a Patriot Party supporter, said Facebook's action against the group was unfair.

PHOTO: PAUL WEAVER/SOPA IMAGES/ZUMA PRESS

Mr. Schwetz said Facebook's action against Patriot Party groups made it more difficult for the movement to grow. He said that was unfair. "Facebook opened itself up to say it's a public platform," he said, but the action against the Patriot Party shows that is untrue.



Source: Internal report titled, 'Information Corridors: A Brief Introduction'

Internal reviews of Facebook's performance around the election's aftermath pointed to the company's inability to keep pace with the speed of its own platform or separate out skepticism of the voting results from incitements to violence.

First reported by BuzzFeed News, one internal report acknowledged that Facebook had been unable to reliably catch harmful content that goes viral, with

the broader Stop the Steal movement "seeping through the cracks" of Facebook's enforcement systems.

The report was optimistic on one front: The company was able to suppress the growth of the Patriot Party.

In the weeks following the Capitol riot, Facebook began studying a data set of more than 700,000 Stop the Steal supporters, mapping out the way information traveled through them.

Armed with the knowledge from that and similar research, Facebook hoped it would be able to sabotage future "harmful topic communities" and redirect what it called "susceptible users" toward innocuous content.

Facebook's strategy for dealing with movements it considers harmful are outlined in a series of internal documents from early this year, from a multidisciplinary group within the company called the Disaggregating Harmful Networks Taskforce.

Under Facebook policies, "an individual can question election results. But when it's amplified by a movement, it can damage democracy," said an April update from the task force. "There is harm in the way movements shift norms and an understanding of collective truth."

Facebook scientists studied how such networks rise, and identified "information corridors"—networks of accounts, pages and groups—that create, distribute and amplify potentially harmful content. The networks span hundreds of thousands of users. Mapping them required artificial intelligence to identify users "most at risk of being pulled into the problematic community," according to one document.

Once a dangerous information corridor is identified, the documents show, Facebook can undermine it. A movement's leaders can be removed, or key amplifiers hit with strict limits on transmitting information.

Unless Facebook chose to disclose such coordinated action, users who weren't themselves removed would never know of the company's interventions.

Facebook deemed the Patriot Party experiment successful enough that the company decided to keep honing its ability to target individual groups, such as the German movement called Querdenken.

Though it shares some affinities with QAnon and anti-Semitism, Querdenken as a group isn't chronically in violation of Facebook policies, according to the company. Officially it preaches nonviolence. But the group had been placed under surveillance by German intelligence after protests it organized repeatedly "resulted in violence and injuries to the police," an internal presentation to Facebook's public-policy team stated.

In April, some of the same employees who tamped down the Patriot Party got to work on an experiment to see if they could suppress Querdenken by depriving it of new recruits and minimizing connections between its existing members.

The Facebook senior security official said the company had acted against Querdenken accounts under a new "coordinated social harm" policy.

Facebook demoted Querdenken content in news feeds and prevented many users from receiving notifications about its posts.

"This could be a good case study to inform how we tackle these problems in the future," one researcher wrote in a document describing the experiment.

—Design by Andrew Levinson. A color filter has been used on photos.

Write to Jeff Horwitz at Jeff.Horwitz@wsj.com and Justin Scheck at justin.scheck@wsj.com

Corrections & Amplifications

"There's a growing set of research showing that some viral channels are used for bad more than they are used for good," Facebook employee Kang-Xing Jin wrote in a note. An earlier version of this article dropped the word "some" from the quotation. (Corrected on Oct. 22)

Appeared in the October 23, 2021, print edition as 'Facebook Increasingly Acts To Squelch Political Groups.'

Copyright © 2022 Dow Jones & Company, Inc. All Rights Reserved

This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to your colleagues, clients or customers visit https://www.djreprints.com.