Internal Alarm, Public Shrugs: Facebook's Employees Dissect Its Election Role

Company documents show that the social network's employees repeatedly raised red flags about the spread of misinformation and conspiracies before and after the contested November vote.

Credit...Illustration by Mel Haasch; Photograph by Anna Moneymaker for The New York Times

Ryan Mac and Sheera Frenkel, *The New York Times* Online, October 23, 2021.

Sixteen months before last November's presidential election, a researcher at Facebook described an alarming development. She was getting content about the conspiracy theory QAnon within a week of opening an experimental account, she wrote in an internal report.

On Nov. 5, two days after the election, another Facebook employee posted a message alerting colleagues that comments with "combustible election misinformation" were visible below many posts.

Four days after that, a company data scientist wrote in a note to his co-workers that 10 percent of all U.S. views of political material — a startlingly high figure — were of posts that alleged the vote was fraudulent.

In each case, Facebook's employees sounded an alarm about misinformation and inflammatory content on the platform and urged action — but the company failed or struggled to address the issues. The internal dispatches were among a set of Facebook documents obtained by The New York Times that give new insight into what happened inside the social network before and after the November election, when the company was caught flat-footed as users weaponized its platform to spread lies about the vote.

From the Document: Content Policy

WHAT HAPPENED
1. From Wednesday through Saturday **there was a lot of content circulating which implied fraud in the election, at around 10% of all civic content** and 1-2% of all US VPVs. There was also a fringe of incitement to violence.
2. There were dozens of employees monitoring this, and FB launched ~15 measures prior to the election, and another ~15 in the days afterwards. Most of the measures made existings processes more aggressive: e.g. by lowering thresholds, by making penalties more severe, or expanding eligibility for existing measures. Some measures were qualitative: reclassifying certain types of content as violating, which had not been before.
3. I would guess these measures reduced prevalence of violating content by at least 2X. However they had collateral damage (removing and demoting non-violating content), and the episode caused noticeable resentment by Republican Facebook users who feel they are being unfairly targeted.

Facebook has publicly blamed the proliferation of election falsehoods on former President Donald J. Trump and other social platforms. In mid-January, Sheryl Sandberg, Facebook's chief operating officer, said the Jan. 6 riot at the Capitol was "largely organized on platforms that don't have our abilities to stop hate." Mark Zuckerberg, Facebook's chief executive, told lawmakers in March that the company "did our part to secure the integrity of our election."

But the company documents show the degree to which Facebook knew of extremist movements and groups on its site that were trying to polarize American voters before the election. The documents also give new detail on how aware company researchers were after the election of the flow of misinformation that posited votes had been manipulated against Mr. Trump.

What the documents do not offer is a complete picture of decision making inside Facebook. Some internal studies suggested that the company struggled to exert control over the scale of its network and how quickly information spread, while other reports hinted that Facebook was concerned about losing engagement or damaging its reputation.

Yet what was unmistakable was that Facebook's own employees believed the social network could have done more, according to the documents.

"Enforcement was piecemeal," read one internal review in March of Facebook's response to Stop the Steal groups, which contended that the election was rigged against Mr. Trump. The report's authors said they hoped the post-mortem could be a guide for how Facebook could "do this better next time."

Many of the dozens of Facebook documents reviewed by The Times have not been previously reported. Some of the internal reports were initially obtained by Frances Haugen, a former Facebook product manager turned whistle-blower.

Ms. Haugen's documents, some of which The Wall Street Journal has published, have created an outcry among lawmakers and regulators, plunging Facebook into one of its worst public relations crises in years. The disclosures from Ms. Haugen, who plans to appear at a hearing in Britain's Parliament on Monday, have resurfaced questions about what role Facebook played in the events leading up to the Jan. 6 Capitol riot.

A rally in Washington in 2019 for QAnon, the conspiracy theory movement. Facebook's researchers identified in 2019 that content on the movement was problematic.Credit...Tom Brenner for The New York Times

Yaël Eisenstat, a former Facebook employee who oversaw safety and security on global elections ads, said the company's research showed it had tried examining its responsibilities around the 2020 election. But "if none of this effects changes," she said, it was a waste.

"They should be trying to understand if the way they designed the product is the problem," Ms. Eisenstat said of her former employer.

Andy Stone, a Facebook spokesman, said the company was "proud" of the work it did to protect the 2020 election. He said Facebook worked with law enforcement, rolled out safety measures and closely monitored what was on its platform.

"The measures we did need remained in place well into February, and some, like not recommending new, civic or political groups remain in place to this day," he said. "The responsibility for the violence that occurred on Jan. 6 lies with those who attacked our Capitol and those who encouraged them."

A QAnon Journey

For years, Facebook employees warned of the social network's potential to radicalize users, according to the documents.

In July 2019, a company researcher studying polarization made a startling discovery: A test account she had made for a "conservative mom" in North Carolina received conspiracy theory content recommendations within a week of joining the social network.

The internal research, titled "Carol's Journey to QAnon," detailed how the Facebook account for an imaginary woman named Carol Smith had followed pages for Fox News and Sinclair Broadcasting. Within days, Facebook had recommended pages and groups related to QAnon, the conspiracy theory that falsely claimed Mr. Trump was facing down a shadowy cabal of Democratic pedophiles.
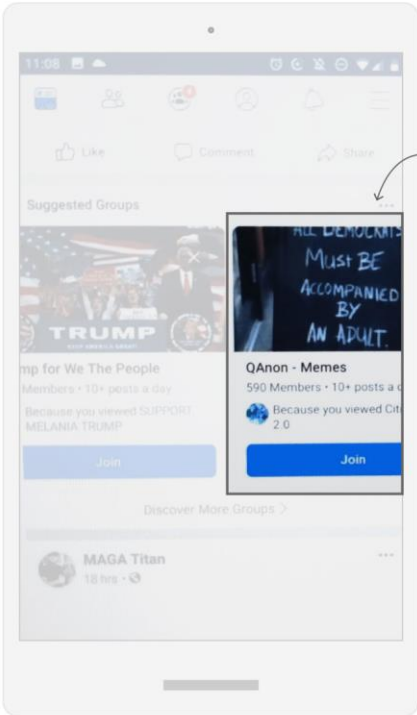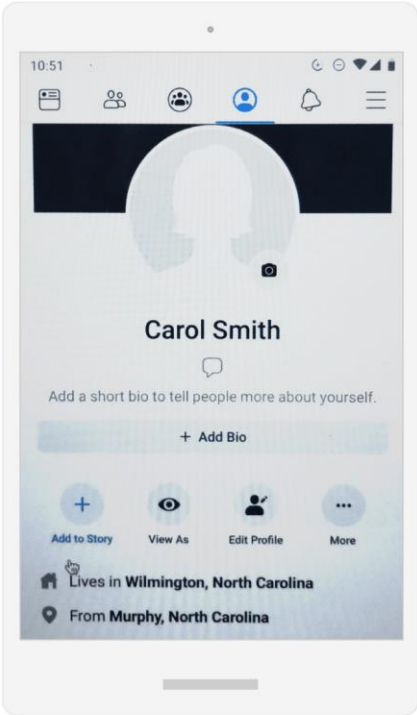
The 'Carol' Account

The account set up for a hypothetical 41-year-old woman named Carol was sparse, with no profile photo and a handful of interests, including parenting, Christianity and civics and community.

Within a week, the account's feed was filled with extreme, conspiratorial and graphic content. Soon after, it began receiving suggestions to join QAnon and similar groups.

By the end of three weeks, Carol Smith's Facebook account feed had devolved further. It "became a constant flow of misleading, polarizing and low-quality content," the researcher wrote.

Facebook's Mr. Stone said of the work: "While this was a study of one hypothetical user, it is a perfect example of research the company does to improve our systems and helped inform our decision to remove QAnon from the platform."

The researcher later ran polarization experiments on a left-leaning test account and found that Facebook's algorithms fed it "low quality" memes and political misinformation. She left the company in August 2020, the same month that Facebook cracked down on QAnon pages and groups.

In her exit note, which was reviewed by The Times and was previously reported by BuzzFeed News, she said Facebook was "knowingly exposing users to risks of integrity harms" and cited the company's slowness in acting on QAnon as a reason for her departure.

"We've known for over a year now that our recommendation systems can very quickly lead users down the path to conspiracy theories and groups," the researcher wrote. "In the meantime, the fringe group/set of beliefs has grown to national prominence with QAnon congressional candidates and QAnon hashtags and groups trending in the mainstream."

Into Election Day

Facebook tried leaving little to chance with the 2020 election.

For months, the company refined emergency measures known as "break glass" plans — such as slowing down the formation of new Facebook groups — in case of a contested result. Facebook also hired tens of thousands of employees to secure the site for the election, consulted with legal and policy experts and expanded partnerships with fact-checking organizations.

In a September 2020 public post, Mr. Zuckerberg wrote that his company had "a responsibility to protect our democracy." He highlighted a voter registration campaign that Facebook had funded and laid out steps the company had taken — such as removing voter misinformation and blocking political ads — to "reduce the chances of violence and unrest."

Many measures appeared to help. Election Day came and went without major hitches at Facebook.

But after the vote counts showed a tight race between Mr. Trump and Joseph R. Biden Jr., then the Democratic presidential candidate, Mr. Trump posted in the early hours of Nov. 4 on Facebook and Twitter: "They are trying to STEAL the Election."

The internal documents show that users had found ways on Facebook to undermine confidence in the vote.

On Nov. 5, one Facebook employee posted a message to an internal online group called "News Feed Feedback." In his note, he told colleagues that voting misinformation was conspicuous in the comments section of posts. Even worse, the employee said, comments with the most incendiary election misinformation were being amplified to appear at the top of comment threads, spreading inaccurate information.

A Stop the Steal protester outside the Capitol in January. Facebook enacted security and safety measures ahead of the election.Credit...Stefani Reynolds for The New York Times

Then on Nov. 9, a Facebook data scientist told several colleagues in an internal post that the amount of content on the social network casting doubt on the election's results had spiked. As much as one out of every 50 views on Facebook in the United States, or 10 percent of all views of political material, was of content declaring the vote fraudulent, the researcher wrote.

"There was also a fringe of incitement to violence," he wrote in the post.

Even so, Facebook began relaxing its emergency steps in November, three former employees said. The critical postelection period appeared to have passed and the company was concerned that some pre-election measures, such as reducing the reach of fringe right-wing pages, would lead to user complaints, they said.

Jan. 6

On the morning of Jan. 6, with protesters gathered near the U.S. Capitol building in Washington, some Facebook employees turned to a spreadsheet. There, they began cataloging the measures that the company was taking against election misinformation and inflammatory content on its platform.

User complaints about posts that incited violence had soared that morning, according to data in the spreadsheet.

Over the course of that day, as a mob stormed the Capitol, the employees updated the spreadsheet with actions that were being taken, one worker involved in the effort said. Of the dozens of steps that Facebook employees recommended, some — such as allowing company engineers to mass-delete posts that were being reported for pushing violence — were implemented.

But other measures, such as preventing groups from changing their names to terms such as Stop the Steal, were not fully implemented because of last-minute technology glitches, according to the spreadsheet.

Mr. Zuckerberg and [Mike Schroepfer](#), Facebook's chief technology officer, posted notes internally about their sadness over the Capitol riot. But some Facebook employees responded angrily, according to message threads viewed by The Times.

A rioter outside the Senate chamber in the Capitol on Jan. 6. Employees kept a running tally of measures that Facebook deployed that day to prevent calls for violence.Credit...Erin Schaff/The New York Times

"I wish I felt otherwise, but it's simply not enough to say that we're adapting, because we should have adapted already long ago," one employee wrote. "There were dozens of Stop the Steal groups active up until yesterday, and I doubt they minced words about their intentions."

Another wrote: "I've always felt that on the balance my work has been meaningful and helpful to the world at large. But, honestly, this is a really dark day for me here."

In a Jan. 7 report, the scope of what had occurred on Facebook became clear. User reports of content that potentially violated the company's policies were seven times the amount as previous weeks, the report said. Several of the most reported posts, researchers found, "suggested the overthrow of the government" or "voiced support for the violence."

Post-Mortems

In March, Facebook researchers published two internal reports assessing the company's role in social movements that pushed the election fraud lies.

In one, a group of employees said Facebook had exhibited "the pattern." That involved the company initially taking "limited or no action" against QAnon and election delegitimization movements, only to act and remove that content once they had already gained traction. The document was earlier reported by The Wall Street Journal.

Part of the problem, the employees wrote, was Facebook's election misinformation rules left too many gray areas. As a result, posts that "could be construed as reasonable doubts about election processes" were not removed because they did not violate the letter of those rules.

Those posts then created an environment that contributed to social instability, the report said.

From the Document: 'Harmful Non-Violating Narratives' Is a Problem Archetype in Need of Novel Solutions

"Retrospectively, external sources have told us that the on-platform experiences on this narrative may have had **substantial negative impacts including contributing materially to the capital riot** and potentially reducing collective civic engagement and social cohesion in the years to come."

Another report, titled "Stop the Steal and Patriot Party: The Growth and Mitigation of an Adversarial Harmful Movement," laid out how people had exploited Facebook's groups feature to rapidly form election delegitimization communities on the site before Jan. 6.

Some organizers sent hundreds of invitations to build the groups, essentially spamming people, the report found. They also asked everyone who joined to invite as many other people as possible, making the groups balloon in size. (Facebook has since begun more closely monitoring the amount of group invites.)

Some organizers of Stop the Steal groups on Facebook also appeared to be cooperating with each other for "growing the movement," the report said.

"Hindsight being 20/20 makes it all the more important to look back, to learn what we can about the growth of the election delegitimizing movements that grew, spread conspiracy, and helped incite the Capitol insurrection," the report said.

Mike Isaac, Davey Alba and Cecilia Kang contributed reporting.
Facebook and the 2020 Election

Ryan Mac is a technology reporter focused on corporate accountability across the global tech industry. He won a 2020 George Polk award for his coverage of Facebook and is based in Los Angeles. @RMac18

Sheera Frenkel is a prize-winning technology reporter based in San Francisco. In 2021, she and Cecilia Kang published, "An Ugly Truth: Inside Facebook's Battle for Domination." @sheeraf