

Facebook, content moderation & disinformation

CNECT I4/F2

FACEBOOK AND THE CODE OF PRACTICE ON DISINFORMATION

Facebook is a signatory of the Code of Practice on Disinformation since its publication in September 2018. Facebook has subscribed to all the commitments of the original Code.

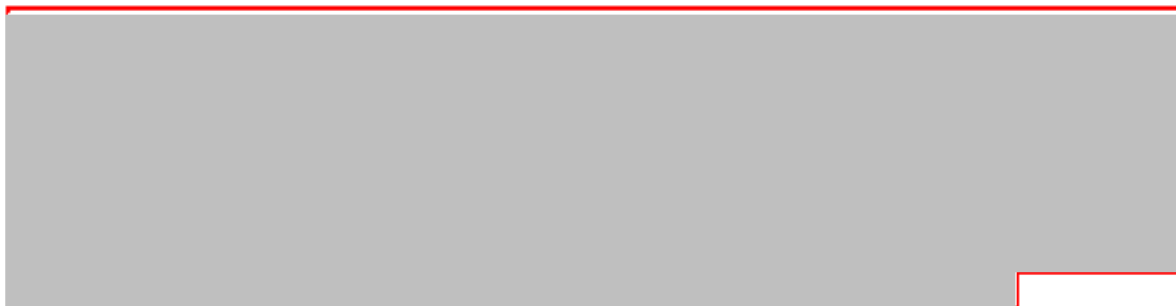
As a signatory of the Code of Practice, Facebook has taken part in the monitoring and reporting programme set up by the 10 June 2020 Communication on COVID-19 disinformation, reporting on measures taken to respond to COVID-19 and COVID-19 vaccines related disinformation.

Art 4.3

A large rectangular area of the document is redacted with a grey fill. The redaction covers several lines of text and a bulleted list. The list contains three items, each followed by a redacted block of text. The first two items are preceded by a single bullet point, and the third item is preceded by a bullet point and a small square icon.

RECENT PRESS COVERAGE AND RESEARCH

Art 4.2
first
indent

A large rectangular area of the document is redacted with a grey fill. The redaction covers the entire content area under the 'RECENT PRESS COVERAGE AND RESEARCH' heading.

Successive news and civil society reports issued only over the past few weeks show allegations of:

- Selective application of national laws (NetzDG). Reset and HateAid performed a test based on 100 hate comments on the platform notably published by far-right politicians. The platform removed 67 of them, applying the terms and conditions of the company, not the legal standard. Where the experiment focused on reporting content violating terms and conditions, the company is reported to have reacted to 50% of the cases.
- The report also shows concerning ad targeting practices, such as ads for conversion therapies displayed particularly to LGBTQI youth or racists and defamatory ads.
- Severe inconsistencies and arbitrariness of policies (e.g. XCheck programme white-listing accounts without prior public information and leading to arbitrary content moderation decisions – revealed by the WSJ),
- Lack of consideration for societal concerns in the design, testing and functioning of recommender systems (e.g. WSJ reports on a 2018 roll-out of new recommender systems maximising engagement and, at the same time, amplifying sensationalist content and misinformation),

Art 4.3

[REDACTED]

[REDACTED]

[REDACTED]

- [REDACTED]
- [REDACTED]
- [REDACTED]

[REDACTED]

Art 4.3

- [REDACTED]

[REDACTED]

Art 4.3

- [REDACTED]

BACKGROUND ON PRESS COVERAGE

On Covid-19 vaccine disinformation

- A recent US study also indicates that people who get most of their news through Facebook may be less likely to be vaccinated against COVID-19⁴.
- However, when in August Facebook shared a report on its most-viewed content on the platform⁵, **none of the top 10 content was divisive political content**. The Washington Post qualifies *“the new report as part of a broader push by Facebook to block or discredit independent research about harmful content on its platform, offering its own carefully selected data and statistics instead”*⁶.

Transparency and working with researchers

- In early August 2021, Facebook cut access to New York University researchers access to ad data⁷, citing privacy concerns. The project's researchers have regularly briefed staffers and lawmakers in the House and Senate and officials at agencies, and the sudden move draw criticism even from the US Federal trade Commission.
- At the end of August 2021, researchers at New York University and France's Université Grenoble Alpes found that from August 2020 to January 2021, articles from known purveyors of misinformation received six times as many likes, shares, and interactions as legitimate news articles⁸.
- At the same time, reports emerged that many Facebook posts from the days before and after the January 6 Capitol Hill riots in Washington are missing from Crowdtangle⁹. After the Capitol Hill riots, the academics said they had planned to analyse what type of content Facebook had removed related to the insurrection to meet its content moderation policies, but they discovered that up to 30 percent of the posts collected around the January 6 riots were missing compared to the original Crowdtangle database.
- This comes at a difficult time, where reports emerged that Facebook is dismantling the team behind the Crowdtangle tool¹⁰.
- On 3 March 2020, AlgorithmWatch launched a project to monitor Instagram's newsfeed algorithm. Volunteers could install a browser add-on that scraped their

4 <https://osf.io/uvqbs/>

5 <https://transparency.fb.com/data/widely-viewed-content-report/> Facebook opted to share four lists that included eclectic mixes of domains and content, with viral animal memes, cooking pages, and sites hawking Christian merchandise crowding out big-name media publishers. Notably, its rankings were drawn from the three-month period between April 1 and June 30, 2021, so they offer no insight into what was popular in the months preceding or immediately following the 2020 election or the Jan. 6

6 <https://www.washingtonpost.com/technology/2021/08/18/facebook-most-viewed-content-report/>

7 <https://www.washingtonpost.com/politics/2021/08/05/technology-202-facebook-shuttered-crucial-tool-oversight-lawmakers-say-it-just-made-their-jobs-harder/>

8 <https://www.washingtonpost.com/technology/2021/09/03/facebook-misinformation-nyu-study/>

9 <https://www.politico.eu/article/facebook-crowdtangle-data-january-6-capitol-hill-riots-transparency/>

10 <https://www.politico.com/news/2020/10/26/censorship-conservatives-social-media-432643> “After an internal battle, the company is currently dismantling the Crowdtangle team after researchers and journalists used the tool repeatedly to trace how far-right, extremist and false content circulated widely across both Facebook and Instagram.”

Instagram newsfeeds. In July 13, AlgorithmWatch ended the project in fear of being sued by Facebook for alleged Terms of Service breaches¹¹.

Political content & manipulation of the service

- Facebook in February 2021 published news that it would expand a trial to reduce the distribution of political content on feeds¹². The company announced it would expand the trial to scale back breaking news and political content in and beyond the United States after receiving “positive feedback” on the changes.
- Critics from the marketing sector said the changes could reduce traffic to news publishers or accounts that post too much political content, based on negative user feedback.
- Still, Facebook's report on the popular content on its platform goes against evidence that stories published by *The Daily Wire*, founded by conservative commenter Ben Shapiro, received more likes, shares and comments than any other news publisher over the past year by a wide margin, according to an NPR analysis of social media data¹³. And Shapiro's personal Facebook page had more followers than *The Washington Post*.
- In September 2021, it was uncovered that troll farms reached 140 million Americans a month on Facebook before 2020 election. Facebook's most popular pages for Christian and Black American content were being run by Eastern European troll farms. Their content was reaching 140 million US users per month—75% of whom had never followed any of the pages. They were seeing the content because Facebook's content-recommendation system had pushed it into their news feeds¹⁴.

CONTACTS: [REDACTED] (CNECT I4)
[REDACTED] (CNECT F2)
[REDACTED] (CNECT F2)

Art 4.1b

11 <https://algorithmwatch.org/en/instagram-research-shut-down-by-facebook/>

12 <https://about.fb.com/news/2021/02/reducing-political-content-in-news-feed/>

13 <https://www.npr.org/2021/07/19/1013793067/outrage-as-a-business-model-how-ben-shapiro-is-using-facebook-to-build-an-empire>

14 Troll farms reached 140 million Americans a month on Facebook before 2020 election | MIT Technology Review