

The Big Read Meta Platforms

Investigating Facebook: a fractious relationship with academia

A growing number of researchers analysing the impact of the social network say their work is being stifled

Madhumita Murgia and **Cristina Criddle** in London and **Hannah Murphy** in San Francisco DECEMBER 6 2021

Last March, Orestis Papakyriakopoulos, a researcher at Princeton University, applied to use a special data access tool that allows academics to do research on Facebook. His goal was to investigate political campaigning on the social network.

The data set contained information on ads related to elections, how they were distributed, to whom and at what cost. But Papakyriakopoulos withdrew his application when he saw what he viewed to be draconian controls on access written into the contract, which he was required to sign.

“Facebook will have the opportunity to review drafts . . . sufficiently ahead of the planned publication or disclosure date . . . solely to identify any Confidential Information or any Personal Data that may be included or revealed in those materials and which need to be removed prior to publication or disclosure,” according to a copy of the draft contract, seen by the Financial Times.

Papakyriakopoulos sought clarification on what constituted “confidential information” but didn’t hear back.

“We could not just start the project and have somebody telling us suddenly that we couldn’t publish,” says Papakyriakopoulos. “[Facebook] said this contract is non-

negotiable because it is mandated by regulators after the Cambridge Analytica scandal.

“That’s just a general excuse that Facebook was using,” he believes.

Earlier this month, the company tried to assuage some of these concerns by launching an updated tool, known as the Researcher API from Facebook’s Open Research and Transparency team (Fort). The tool is currently available to an unpublished list of two dozen research institutions invited by Meta, Facebook’s newly named parent company.



Orestis Papakyriakopoulos withdrew his application to Facebook after finding draconian controls on access written into his contract, which he was required to sign

Yet the incident is one of a multitude of examples of Meta’s uneasy relationship with researchers who are seeking to understand the potentially harmful social effects of the platform. A growing number of academics complain that the company puts up

excessive roadblocks or tries to stifle research that might cast it in a negative light.

Meta's relationship with its own researchers was tested recently after [whistleblower Frances Haugen](#) leaked troves of documents by internal company researchers but whose conclusions had been buried. The potential harms from the company's platforms that they outlined range from election misinformation on Facebook to Instagram posts that exacerbate mental health issues amongst teenage girls concerned about body image.

Her revelations have fed a narrative that Facebook operates on a growth-at-all-costs mentality, despite the growing criticism of the impact of the social network on society and politics.

Several external and independent academics and researchers told the FT that they are concerned the company's stranglehold on its data is a matter of public concern. Some have even compared it to the way that industries such as tobacco have in the past attempted to shape and manipulate academic research.

"Facebook is trying to block research on its platform quite systematically, and that goes against the principles of academia and the public interest. It's closest to what Big Tobacco was doing . . . setting up research institutes and commissioning research that isn't really research," says a researcher who has worked on a Meta-funded research project and requested anonymity to prevent professional backlash. "The number of cases is now building up to a picture of a consistent war on independent academia."

A Meta spokesperson said it partners with academic researchers and is building products that support their work. "New technologies and partnerships mean that we can share data sets in a privacy-protective environment that significantly advance academic research. Like the rest of the field, we continue learning about the best ways to share this data while preserving the privacy of the people who use our services," the company said.

For some observers, the Haugen revelations and the complaints of academic researchers both point to the same issue — an absence in the public realm of any real understanding of how the algorithms on the social media platform work, which allows the company to fend off criticisms about any known negative impacts of its technology.

"The biggest revelation from whistleblower Frances Haugen's documents is what Facebook is able to hide — and that applies to all the Big Tech companies; they are completely opaque," says Emma Briant, a propaganda and influence operations researcher at American University in Washington, DC and Bard College. "Our entire

researcher at American University in Washington DC and Bard College. Our entire democracy then rests on the moral choices and bravery of individuals within the company who come forward, or don't. There is a really disturbing pattern emerging here.”



Whistleblower Frances Haugen leaked multiple documents outlining some of the harms Facebook can cause © Geert Vanden Wijngaert/AP

The Cambridge Analytica hangover

There are good reasons for Meta to be very careful about the way it manages academic scrutiny of its data

SCRUMBY OF ITS DATA.

The company became much more restrictive to outsiders after the Cambridge Analytica scandal in 2018, when it emerged that a small political consultancy obtained the personal data of some 87m Facebook users via a third party without proper consent. In 2019, the company paid a \$5bn settlement to the US Federal Trade Commission over these privacy violations. Since then, Meta has walked a tightrope, trying to balance users' privacy with more transparency, all the while trying to maximise the growth of its platform.

Briant, who has been studying Cambridge Analytica for more than a decade, says if access to the company's data is not handled well, that information could become available to governments and other actors. "A multitude of researchers are seeking access, not all of whom would have a strict university ethics process, transparent purposes and assured security," she says.

The company says that its user data is governed by privacy laws across the world, such as Europe's General Data Protection Regulation, and it therefore needs to manage carefully any access it gives to third parties.

In the case of the new tool it has developed for academics, Meta says that researchers do not need to sign this contract as the company wants to make it less taxing to access public data and has taken on feedback from researchers who are given early access.

The company says it hopes to open up access to the tool more widely from February to vetted researchers, who prove that they are affiliated with a university and undergo some training in how to use the system.

However, it says it does still require pre-publication review when research involves sensitive user data. The company also said it would never ask academics to modify their findings, but it would highlight proprietary or identifying information that needs to be removed.

The documents released by Haugen suggest Meta had withheld its own internal research on potential ill effects from the public. Some of those in-house researchers have long complained about overly restrictive contracts. Documents seen by the FT show that academics who are contracted to work internally for the company, either as temporary or permanent staff, are held to the same restrictions as non-academic staff.

In at least some contracts, the company claims ownership of all "inventions" —

including blog posts, books and future research papers — that use any information or knowledge obtained during the course of working there. For a year after an academic's employment, all new work (even if it does not relate to the social network) must be declared to Meta, explicitly detailing why the company cannot claim it as its own.

Academics who have worked at Meta told the FT they felt muzzled by such contracts. The researchers were concerned their future academic work and publications would be affected if they were restricted in their use their insights and experiences at the company — one of the main reasons somebody may go to work for Meta in the first place.

Meta said it hires academics to work internally because of their expertise, but that they try to explain the parameters of the relationship very clearly. It acknowledged confidentiality clauses were a source of contention but said the contracts were drawn up by its legal team.



Facebook's fake news 'War Room' is designed to assuage public concern about the fake accounts, misinformation and foreign interference on its site during elections © David Paul Morris/Bloomberg

“Any contract that involves access to data involves confidentiality clauses and this is no exception,” a Meta spokeswoman said. “We use standard confidentiality language with numerous carve-outs for specific situations, and do not have a non-compete

clause that restricts the future work of academics who work with us.”

One contractor who worked on the Facebook AI research team but turned down a full-time job said: “Facebook out of all the Big Tech companies is the least attractive to [academic] talent.”

Growing complaints from independent academics

Papakyriakopoulos is far from the only researcher to balk at conditions imposed by Meta.

In August, the company deactivated access to its platforms for two researchers at New York University, claiming they had breached its guidelines. But the researchers accused it of trying to shut down their work as it revealed the company was amplifying partisan misinformation in the ads it promoted.

“Facebook has not shown itself to be a good partner here,” says Laura Edelson, the lead researcher involved. “If you look at what they’ve done with their own internal research . . . that would not have seen the light of day were it not for other events. I think that’s an attitude to research that makes a lot of independent researchers pretty wary.

“[Previously Facebook] opened all the windows and people looked in, and now we don’t like what we are seeing, the reaction has not been to clean up the house, it has been to close the windows.”

The company has also been accused of interfering in the work of independent researchers they fund. In 2020, Facebook donated \$1m to the Partnership for Countering Influence Operations, a research project within the Carnegie Endowment for International Peace, a non-partisan think-tank in Washington DC. The goal was to facilitate independent investigation into the effects of online manipulation and misinformation.

While it started as a genuine research project, Meta’s influence allegedly increased with time. “It gradually became more and more directly steered by Facebook,” says a researcher close to PCIO. “It became daily instructions filtering through, messages saying they had heard something . . . or seen a paper they didn’t like. It was subtle messages from Facebook, always through other people.” Original investigations were discouraged, the person said, and their output became mostly summaries of the current literature. Meta said it did not interfere with the work of studies it funds

current literature. Meta said it did not interfere with the work of studies it funds.

Rebekah Tromble, a professor at George Washington University who studies the spread of misinformation online, says the company has used the GDPR, Europe's privacy laws, as an excuse to prevent access to data that researchers request.

Tromble was one of the original members of Social Science One, a non-profit initiative founded by Harvard and Stanford professors in 2018 aiming to be a data broker between Facebook and academics. The first data set on offer included "almost all" public links shared and clicked by Facebook users globally, around a petabyte of data.

"One of the things that was profoundly concerning for academic researchers and social scientists, in particular, is if we want to understand cause and effect, we have to be able to look at data at the individual user level," says Tromble. "But Facebook was just saying no and using GDPR as the key barrier for them to do this."



An anti-Facebook protest outside parliament. Frances Haugen said both to UK and European parliaments that Facebook is 'very good at dancing with data' © Tolga Akmen/AFP via Getty Images

Tromble approached policymakers in Brussels to clarify and discovered that GDPR had a special exception specifically for academics to be able to access data. Meta said these exceptions did exist, but there was a lack of clarity around whether this was

applicable to the company.

Eventually, the project was undermined when Facebook handed over what some of the researchers claimed to be incomplete data to researchers, excluding around half of US users, rendering months of work and analysis unusable.

Facebook said there was an error in the data set that impacted some research but it has worked hard to update the data since the incident.

“The problem is that as long as Facebook and other platforms completely control what they are willing to share with researchers and that data can’t be independently verified in any way, we are always vulnerable to the critique that we don’t know for sure that our analyses are right,” Tromble says. “What we see is that platforms actually use this against us.”

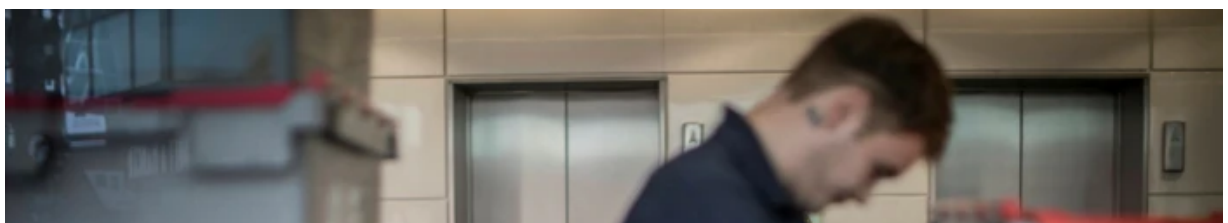
Social media and democracy

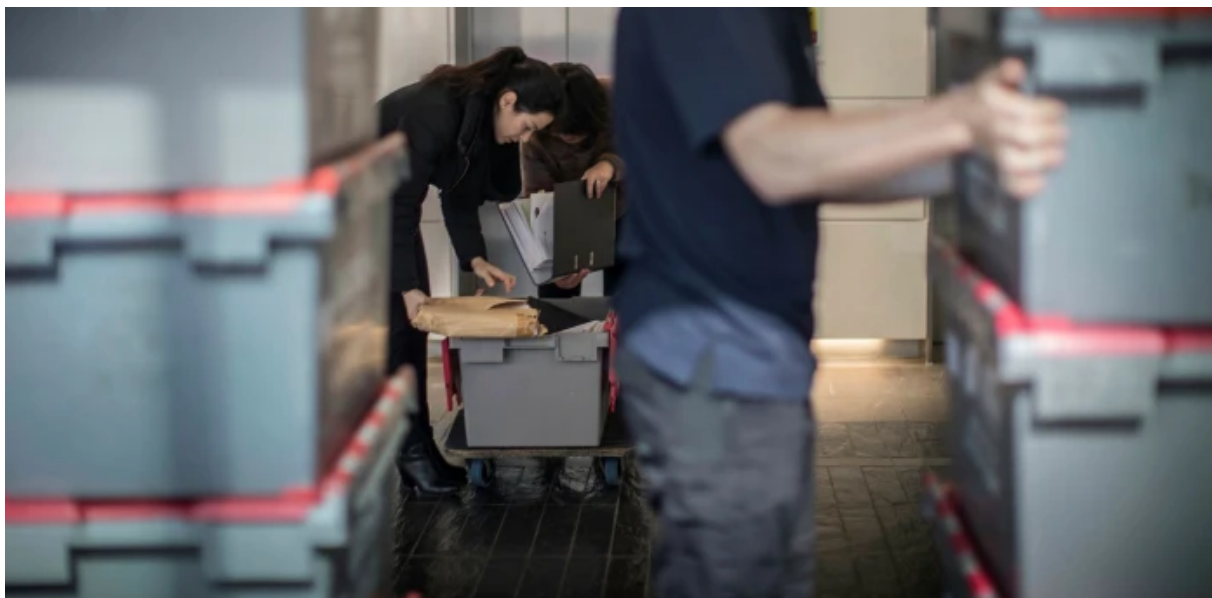
Meta says its work with academics has been evolving over the past three years and, in March, it set up a dedicated academic partnerships team to act as an internal liaison for researchers who want to conduct studies on Facebook or Instagram.

Currently, a team of 17 external academics are working with Meta on a new project, known as Election 2020, a series of studies into the role of social media in democracy today. Those involved hope it can provide a model for future collaboration with the company.

To protect academic independence, the researchers do not receive money from Meta (although it does fund parts of the research), Meta cannot review work prior to publication, an independent academic observer is overseeing the research process, and participants must opt in to the research of individual-level data. Researchers the FT spoke to said the project was going well, with little pressure or interference thus far.

However, to protect user identities, in some cases the researchers cannot access data directly and have to depend on Meta to mine it on their behalf.





Cambridge Analytica's offices, 2018. Facebook became much more restrictive to outsiders after the scandal that involved the company © Charlie Bibby/FT

“I was a little wary of entering into a research partnership with Facebook,” one researcher involved in the project says. “I haven’t felt any pressure . . . [But] it is a more cumbersome process, it is not what I am used to.”

Despite early positive signs on the Election 2020 project, some researchers who are part of it still feel that the power lies squarely with Meta, which can choose what data to share and how. They believe that laws requiring companies to provide data and information for public benefit is crucial for academics to be able to conduct truly independent research on social media platforms.

Tromble, who is one of the researchers on the Election 2020 project, says: “I very firmly believe that without regulation to mandate access, we simply won’t be able to get the type of transparency and accountability that we all desire.”

Frances Haugen echoed this when speaking to the European Parliament about the Digital Services Act (DSA), a proposed bill that clarifies the responsibilities of Big Tech companies in Europe. Haugen urged lawmakers ‘to encourage the platform to let information stored on it be widely available, rather than just for “vetted academics”, as recommended in the current proposal.

Haugen told both the UK and European parliaments that Facebook is “very good at dancing with data”, and said legislation should compel Facebook to explain the information it hands over, including queries used to pull the data.

European digital rights campaigners such as AlgorithmWatch are also campaigning for tougher regulation to compel data access. The DSA currently only includes this requirement for university academics. “We think it is crucial but that it should be amended to include not just researchers with academic affiliations but also those from civil society and journalists,” says Angela Müller, who leads the policy and advocacy team at Algorithm Watch.

In the US, academics have drafted a bill that would allow the FTC to set mandatory data and information sharing requirements for social media platforms, with penalties for researchers and companies that violate the requirements. Two senators, one from each party, are now planning to introduce legislation building on this proposal.

“We need to have some way of imposing a cost if social media companies don’t do this — otherwise it’s just a risk,” says Nate Persily, a professor at Stanford Law School who drafted the bill, and co-led the Social Sciences One project. “My view is we’ve got to get this done right away before the 2024 election.”

[Copyright](#) The Financial Times Limited 2021. All rights reserved.
