

- An October 2023 report from the Reuters Institute of Journalism, University of Oxford

Here's what our research says about news audiences on Twitter, the platform now known as X

According to our data and despite undergoing many changes, the use of Twitter for news has remained remarkably stable in the last decade

<https://reutersinstitute.politics.ox.ac.uk/news/heres-what-our-research-says-about-news-audiences-twitter-platform-now-known-x>

- A very thorough Study by the Pew Research Center on the use of TWITTER

<https://www.pewresearch.org/journalism/2021/11/15/news-on-twitter-consumed-by-most-users-and-trusted-by-many/>

- On Trump and his use of twitter

<https://theconversation.com/i-analyzed-all-of-trumps-tweets-to-find-out-what-he-was-really-saying-154532>

https://www.lemonde.fr/les-decodeurs/article/2017/01/24/compulsive-et-emotionnelle-quelle-est-l-utilisation-de-twitter-par-donald-trump_5068293_4355770.html

<https://www.nytimes.com/interactive/2021/06/07/technology/trump-social-media-ban.html>

(here is the same article as a [gift](#)... 😊)

- The Long Read: Twitter under Elon Musk's management

<https://www.theguardian.com/technology/2023/jan/29/tears-blunders-and-chaos-inside-elon-musk-twitter>

Document A - Opinion | Twitter Gave Us an Indispensable Real-Time News Platform. X Took It Away.

The Arab Spring cemented Twitter's indispensable role in fast-moving international stories. Elon Musk has destroyed that and given us something useless — and dangerous.

5 Opinion by CLAIRE BERLINSKI

Politico, 11/08/2023

Claire Berlinski is a Paris-based journalist and the editor of the Cosmopolitan Globalist.

I first began using Twitter, now X, after the 2010
10 earthquake in Port-au-Prince, when I could not find my family. I was in Washington, D.C. Within days of the earthquake, hundreds of volunteers, using a crisis-mapping platform called Ushahidi (Swahili for “witness”), began logging calls for help via cellphone
15 texts and Twitter, putting them on a map and sending that information to search and rescue teams. I only

found my family because Facebook and Twitter worked even when their phones and email did not.

Not long after came the Arab Spring, when Twitter
20 was a source of crucial, sometimes lifesaving information and a powerful coordinating tool for demonstrators. For journalists and analysts, it was matchless for watching events unfold in real time. I was in Istanbul during the 2013 Gezi Park
25 demonstrations, which began as a protest against the destruction of a park and turned into a massive uprising against the Turkish government. Like everyone, I used Twitter to figure out which streets were safe. No news agency could replicate this function for citizens on the
30 ground. But I used it, too, to tell the world what was happening in Turkey.

That was the beginning of an era. Twitter was obviously imperfect, but it quickly became *the* central platform for gathering information and networking in politics and international affairs, especially in crises where traditional media coverage was thin.

The war in Gaza marks the end of that era. X CEO Elon Musk has so profoundly undermined the functions that made Twitter useful in an international crisis that it is now counterproductive to turn to it if you hope to understand what's happening on the ground when news breaks around the world. Over the course of 10 years, X has evolved from indispensable to useless.

It's fitting that along with his other changes, Musk changed the name of the company. Literally and figuratively, we're witnessing the first post-Twitter major world conflict.

Disinformation and misinformation proliferated on the platform prior to Musk's takeover. But the difference in degree is so significant now as to amount to a difference in kind. Almost every significant change Musk has made has reduced its value as a source of reliable information from and for people affected by a disaster, and every change, similarly, has increased its utility to malicious propagandists and scammers.

When Musk arrived, he dissolved Twitter's Trust and Safety Council, the advisory group of some 100 independent civil, human rights and other organizations that helped the company to combat misinformation, and fired its full-time Trust and Safety employees, including all but one member of the AI ethics team that monitored algorithmic amplification.

He blew up the old verification system. A blue check once signified that the account belonged to someone whose identity had been confirmed and who fell under one of the following categories: government; companies, brands and organizations; news organizations and journalists; entertainment; sports and gaming; activists and organizers; content creators; and influential individuals.

The prevalence on Twitter of verified journalists, academics, researchers and government sources made it possible, in a crisis, to quickly find reliable people who were on the ground and who could probably be trusted to report what they were seeing in reasonably good faith. Now, the blue checkmark signifies only that the owner has paid for it. When you buy a blue check, your posts go to the top of the search results and replies, irrespective of others' desire to see them.

X now pays users based on the number of views they receive, creating a massive incentive to post sensationalistic and inflammatory lies. On-the-ground witnesses — and worse, *people who need help* — can't

reach their audiences unless they have a costly blue check mark and a willingness to compete with the most outrageous promoted content.

Musk has stripped headlines and summaries off article previews. Images are no longer accompanied by context, making it that much easier to misunderstand them and that much less likely that users will read the article. Meanwhile, Musk promotes — directly, via his own tweets and algorithmically — conspiracy theorists, Russian war propagandists, hostile-state media, foreign and domestic extremists and engagement farmers who exploit pain and tragedy to gain followers.

Musk has created the conditions for a flood of misinformation and fake accounts — a problem that has proven particularly acute during the current war in Gaza. AI-generated images are passed off as real: In a photo shared millions of times, for example, a man is seen carrying his children out of the ruins of a bombed building. The post says, "An image is worth a thousand words." The Palestine flag suggests the image is from Gaza. A BBC fact-checker has confirmed the image is completely fake, generated by AI.

A massive number of accounts posting images that purport to be from Gaza are in fact posting images from unrelated conflicts. These tweets have racked up millions of views and shares.

According to Cyabra, an Israeli analysis firm, pro-Hamas forces have launched a coordinated influence operations campaign involving tens of thousands of fake profiles. As a result, one in five social media accounts participating in the conversation about the war in Gaza are fake. One in four pro-Hamas profiles are fake. It's not clear who is creating and using these fake profiles to spread disinformation, but it could be anyone from Russian internet trolls to antisemites to far-right hucksters who are eager to make a buck.

Accounts that were once clearly labeled as state-affiliated, such as that of Iran's Press TV, are no longer distinguished from others. In September, an EU report found that the "reach and influence" of Kremlin-backed accounts on social media, and on X in particular, had increased in 2023.

In another study, the EU found that disinformation was more easily found on X than on any other social media platform. It also received more engagement than it did on any other platform.

That report found that X had the highest ratio of what the authors called "disinformation actors" to real posts. "The average engagement with mis/disinformation content found on X is 1.977 times as high as the average engagement with non-mis/disinformation," the authors

135 wrote. In other words, X users are twice as likely to engage with lies as the truth.

Snuff videos have proliferated on the platform, too. A pro-Kremlin account shared a video of a beheading in Ukraine in April. Endless videos of 140 atrocities purportedly committed by Hamas or by Israel clatter across my feed: It is all but impossible to use X now without seeing a video that shows, or purports to show, the murder of a human being.

Meanwhile, good sources of information are leaving 145 the platform. Many of the most useful voices are now gone. Reporters have fled, largely moving to Bluesky. But Bluesky can't replace X yet; its network is too small. You can use it to talk to other journalists, not so much to find sources or promote your work to your 150 readers.

To judge by the responses to the fake tweets, most people have no idea they're fake. Musk certainly doesn't. Recently, he recommended (in a since-deleted tweet) that his followers follow two well-known 155 disinformation accounts — one of them, for example, provides such helpful analysis as, "The overwhelming majority of people in the media and banks are zi0nists." When Musk suggests something like this, it is not just his 162 million followers who see it. You can mute him, 160 but unless you do, everything Musk says is now forced into the timeline of every user of the platform, whether or not they follow him.

Some will reply, correctly, that the mainstream media hasn't covered itself in glory during this conflict 165 either. Many news organizations have made serious mistakes in the past weeks. But the difference between sloppy or biased news coverage and millions of *deliberate* lies — none retracted — is significant. If you read the former, you'll still be on this planet. If you 170 read the latter, you'll be in an alternate universe.

As for Israelis or Palestinians hoping to use the platform to exchange lifesaving information? Forget it. For eight bucks, anyone can impersonate the official account of the IDF. (Impersonation is still officially 175 against X's rules, but the lack of a real verification system makes it easier to do for longer.) How could Israelis trust warnings and updates from their government or from each other on X when Hamas or Iranian operatives can so readily pretend to be anyone 180 they please? How would Palestinian rescuers use X to locate people who need help when every image they see directs them to Syria?

This state of affairs is massively deleterious to American national security. Members of Congress are 185 as vulnerable to hostile disinformation as anyone else. One morning, I watched a number of Russian

accounts, including that of former Russian President Dmitry Medvedev, begin simultaneously to push out the line that Israel had been attacked with weapons the U.S. 190 sent to Ukraine, which Israelis immediately denied. By afternoon, U.S. Rep. Marjorie Taylor Greene (R-Ga.) was asserting this as fact.

Organized, state-sponsored information operations aim to generate support in the U.S. for policies that are 195 not in our interest and dampen support for policies that are. "Feel the difference," wrote Medvedev on X recently on a post that showed two photos: on the left, a vast sea of protesters against Israel — a U.S. ally — and, on the right, a pitiful handful of protesters against 200 Russia's actions in Ukraine. That Russian information operations on social media have inflamed the global orgy of antisemitism we've seen since Hamas' attack goes without saying.

X has long been a critical source of information for 205 journalists, and the news cycle now demands reporting at X speed. For most users, the website was a diversion. For journalists, it was essential. We built our audiences and our careers on Twitter. We invested years of our labor and creativity — for free. *We* built it, 210 and *we* made it valuable.

The degeneration of the quality of information on X means that journalists who are still on the platform waste far more time looking for the signal in the noise. They waste more time running down rumors. They are 215 at greater risk of sharing fake information. They are doubtless absorbing narratives and framings from X shaped by disinformation, even if they're not sharing falsehoods. It's far from clear how the media will adapt to a post-Twitter world.

220 At least in the short term, the market won't be able to solve this problem because Twitter's value to consumers was owed to its market dominance. Everyone used it. A number of competitors are now trying to fill the void, but because the microblogging 225 market is now fractured, no company can play the central role Twitter played.

That might be why some governments are stepping in to fix these problems instead. On Oct. 13, the EU launched a probe of X over the spread of disinformation 230 and violent content related to the Israel-Hamas conflict, the first step in an investigation to determine whether the platform is violating the EU's new Digital Services Act.

Congress should do the same. The situation is 235 intolerable to a democracy. Musk is harming American security by gift-wrapping X and handing it over to terrorists and hostile governments who are running information operations against democratic ones. In the

long term, it's finalizing the divorce between the public
240 and reality.

But Congress seems incapable. It was hard enough
for the House to elect a speaker. X has been left to
govern itself, and it has no interest in preserving the
uniquely important role Twitter played by allowing real
245 people, around the world, instantly to find and
communicate with one another during a crisis.

It's fashionable to be cynical about X, but at its height,
it was an astonishing human achievement. It cannot be

rebuilt and it's hard to see how it could be built again.
250 One man destroyed it. It is one of history's great acts of
vandalism.

We still badly need reliable, real-time information
about global conflicts on a platform with the reach that
Twitter once had. But X won't ever again be that
255 platform.

Trump and Twitter

TEXT 1 -How Trump uses Twitter to distract the media – new research

The conversation, 12 novembre 2020

Auteurs : **Ullrich Ecker** Professor of Cognitive Psychology and Australian Research Council Future Fellow, The University of Western Australia, **Michael Jetter**, Associate Professor in Economics, The University of Western Australia, **Stephan Lewandowsky**, Chair of Cognitive Psychology, University of Bristol

In both the lead up to and the immediate aftermath of the US presidential election, President Donald Trump made claims of voter fraud and a rigged election, using all channels available to him, including Twitter. Despite the apparent lack of evidence for these accusations, they have arguably influenced the beliefs of millions of Americans. Twitter has been a primary means by which the president has sought to set the agenda. Since he first took office, many people have speculated that some of Trump's tweets were deployed to distract from negative media coverage. For example, when the press reported on the US\$25m Trump University settlement, he tweeted about the Hamilton play controversy. When COVID-19 failed to "just go away" but instead took a stranglehold on the US, he tweeted about the "OBAMAGATE!" conspiracy theory.

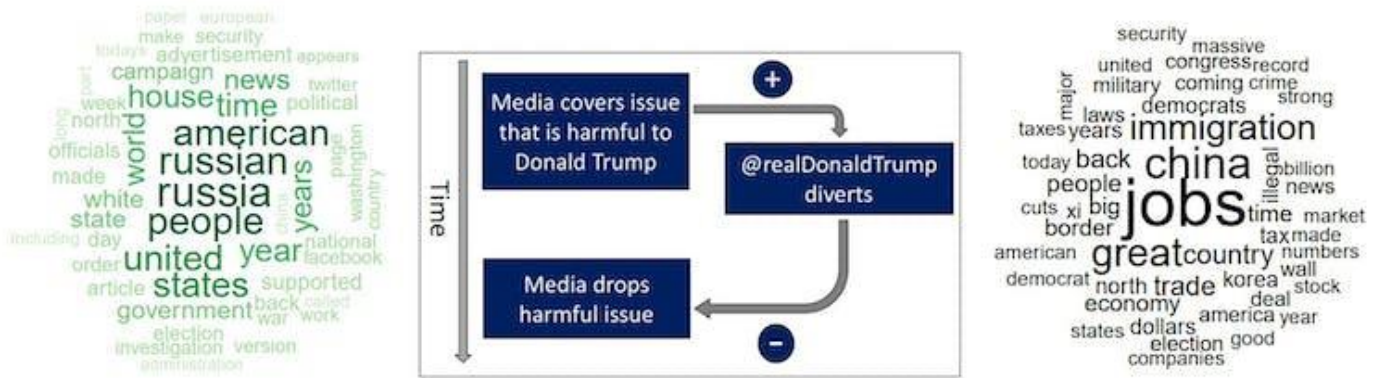
At least some of these distractions seem to have worked. For example, our previous research showed how there was far greater public and media interest in the Hamilton controversy than the Trump University settlement. But the evidence had been anecdotal – until now.

Our new research presents the first empirical evidence that Trump's tweets systematically divert attention away from topics that are potentially harmful to him. Perhaps even more importantly, we found that this diversion works and suppresses subsequent coverage of potentially harmful news stories.

We asked two questions: is potentially harmful media coverage followed by increased diversionary Twitter activity by Trump? And does such diversion reduce subsequent media coverage of that topic?

To test the hypotheses, we focused on the content of the New York Times (NYT) and ABC World News Tonight (ABC) headlines and all of the approximately 5,000 Trump tweets during his first two years in office. We chose the Mueller investigation into potential collusion with Russia as the harmful topic. We then selected a set of keywords – "jobs", "China" and "immigration" – that we assumed would be Trump's go-to topics at the time, based on a systematic content analysis of his campaign materials and major talking points.

The team hypothesised that the more the NYT and ABC reported on the Mueller investigation, the more Trump's tweets would mention jobs, China and immigration, which – if the diversion were successful – would then be followed by less coverage of the Mueller investigation by NYT and ABC the following day. The logic is illustrated in the graphic below.



The word cloud on the left contains the 50 most frequent words from all articles in the NYT. The 50 most frequent words occurring in Trump's tweets are on the right. Author provided

To explore the robustness of these results, we also conducted an expanded analysis that considered the president’s entire Twitter vocabulary as a potential source of diversion. This analysis corroborated our findings: “jobs” and “China” were still Trump’s top picks, but “tax”, “crime” and “North Korea” also featured prominently as diversionary topics.

To illustrate, when we considered “placebo topics”, such as Brexit, no diversion was observed. These placebo topics presented no political threat to Trump and were selected to span a variety of unrelated domains, including football and gardening. In other words, only media reports on Mueller/Russia – but not reports on placebo topics – resulted in an increase in diversionary Trump tweets.

Strategic diversion is not a new political tool. It was the topic of the 1997 film “Wag the Dog”, which saw commentators draw parallels to then President Bill Clinton’s handling of the Monica Lewinsky scandal. However, social media has allowed political leaders more direct and immediate access to their constituents and the media. Our analysis shows that they can use this pathway effectively to divert.

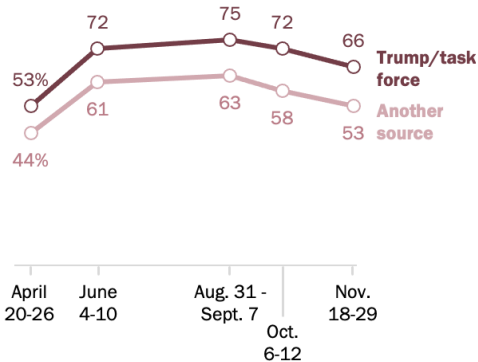
See also:

Causing a rupture with a close ally, endorsing a hate group, undermining a cabinet member ... it's all in a morning's work for the president's restless thumbs

●Republicans who relied on Trump for news in 2020 diverged from others in GOP in views of COVID-19, election
<https://www.pewresearch.org/journalism/2021/02/22/republicans-who-relied-on-trump-for-news-in-2020-diverged-from-others-in-gop-in-views-of-covid-19-election/>

Republicans who turned to Trump for news were more likely to say COVID-19 overblown, media coverage inaccurate

Among Republicans/Lean Republicans who relied most on ____ for coronavirus news, % who said the outbreak had been made a bigger deal than it really is

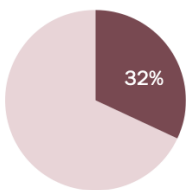


For roughly three-in-ten Republicans, Trump was a major source of election and/or COVID-19 news

% of Republicans/Lean Republicans who ...

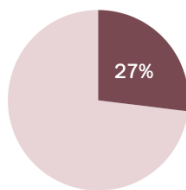
Relied most on Trump and the White House task force for COVID-19 news

APRIL 2020



Said the Trump campaign is a major source of election news

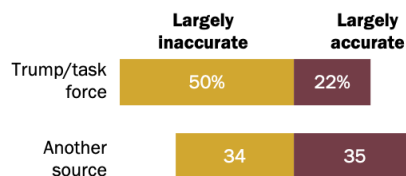
SEPTEMBER 2020



Source: Surveys of U.S. adults conducted April 20-26, 2020, and Aug. 31-Sept. 7, 2020.
 "How Americans Navigated the News in 2020: A Tumultuous Year in Review"

PEW RESEARCH CENTER

Among Republicans/Lean Republicans who relied most on ____ for coronavirus news, % who said in April 2020 that media coverage of the outbreak was ...



Note: In the second graphic, respondents who did not provide an answer not shown, and those who said neither phrase reflects their views not shown.
 Source: Surveys of U.S. adults conducted Nov. 18-29, 2020, and April 20-26, 2020.
 "How Americans Navigated the News in 2020: A Tumultuous Year in Review"

PEW RESEARCH CENTER

TEXT 2 -The nerve center of the American news cycle

AXIOS, Jun 9, 2020 - [Technology](#)

The fast-moving world of Twitter has become the nerve center of the American news cycle — as evidenced by record-breaking downloads and engagement for the service last week.

Why it matters: Twitter is our mediaverse's grand interface between journalism and social media. While news organizations play a central role in sharing links to their coverage on Twitter, much of the visual content shared in real time during breaking news events like protests is shared by everyday users.

- **Thanks to the ubiquity of smartphones**, a camera can be just about everywhere, but no professional news team can be everywhere.
- **The upside** is that news organizations can quickly access footage that helps bring stories to light on TV.
- **The downside** is that these videos, often difficult to verify, are immediately presented with very little context to large viewer numbers online.
- **What users see** is determined by Twitter's social graph and its selection algorithm, rather than by editors or reporters applying (hopefully) professional judgment.

Our thought bubble: Twitter sets the news cycle's pulse because so many journalists are addicted to it. Its power is in agenda-setting. But that's all happening instantaneously and out in the open, not behind the closed doors of an editors' meeting.

The big picture: Twitter and other online platforms have opened a wide path for powerful images — like those of the killing of George Floyd — to reach the public.

- Police protest supercut videos have proven wildly popular online, per The New York Times (see [HERE](#)), helping to spur the #DefundThePolice movement growing alongside the racial protests. Some videos have gotten up to 40 million views on Twitter alone, and millions more on other platforms like YouTube and Instagram.
- A video of a woman calling the police on a black bird watcher in Central Park on Memorial Day also racked up over 40 million views on Twitter.

But the constant flow of sensational content from everyday users, often lacking key context and unverified, also promotes polarization and the spread of misinformation.

- **There's a danger** that for every really important story that comes to light thanks to Twitter, there are others that become big stories even before anyone knows if they're even true.
- **Case in point:** In one story that recently went viral, a man was misidentified on Twitter and other platforms last week as the person who'd been caught on video attacking people for posterizing rally flyers in D.C. Online sleuths had wrongly connected him with the incident thanks to data his bike-riding app publicly recorded.

By the numbers: Wednesday was the number one day in Twitter's history for downloads with 677,000 globally, per app measurement company Apptopia. It also set a record for daily active users on Twitter in the U.S. that day, with 40 million.

Be smart: Twitter has long stood out as the social media network with some of the most news-focused users, per Pew Research Center.

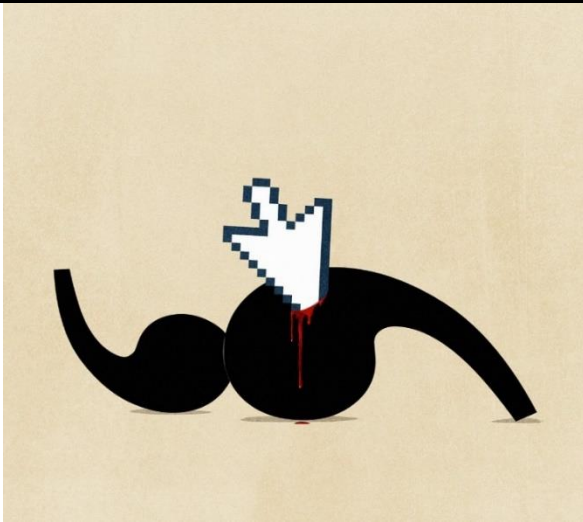
- While videos from one site are often reimagined and then reposted on other platforms, like Instagram and YouTube, often the raw footage from live news events is first posted on Twitter.
- Around seven in ten adult Twitter users in the U.S. (71%) get news on Twitter. About 1 in every 5 U.S. adults uses Twitter.

Twitter's architecture suits it to be a go-to place for news. Its combination of short messages and its option to view a simple chronological feed makes it a good tool for news junkies and journalists alike.

- Because most of the content is public, it's easy to share things broadly quickly, unlike a platform like Facebook, where user-generated news content is shared with friends, spreading slower.
- The platform has also been popularized by world leaders that are super-users, like President Trump. These figures make news when they tweet, and their messages can start global conflicts, disrupt markets or drive civil unrest.

History lesson: The recent protests sweeping the nation, and the world, are similar to previous moments that helped make social media the news engine that it is today.

- Major civilian uprisings, like the Arab Spring, have helped to democratized information in unprecedented ways.
- But today, Twitter and its peers are mired in political culture wars. And their ability to act as neutral platforms for free speech is being tested by world leaders spewing misinformation.



[Leaders](#) | Who controls the conversation - **TEXTE EN VERSION AUDIO** sur CdP

TEXT 3 - **How to deal with free speech on social media**

It is too important to be determined by a handful of tech executives

The Economist, Oct 22nd 2020

It is the biggest antitrust suit in two decades. On October 20th the Department of Justice (DOJ) alleged that Google ties up phone-makers, networks and browsers in deals that make it the default search engine. The department says this harms consumers, who are deprived of alternatives. The arrangement is sustained by Google's dominance of search which, because of a global market share of roughly 90%, generates the advertising profits that pay for the deals. The DoJ has not yet said what remedy it wants, but it could force Google and its parent, Alphabet, to change how they structure their business. Don't hold your breath, though: Google dismisses the suit as nonsense, so the case could drag on for years.

Action against Google may seem far from the storm gathering against Facebook, Twitter and social media. One is laser-focused on a type of corporate contract, the other a category 5 hurricane of popular outrage buffeting unaccountable tech firms for supposedly destroying society. The left says that, from the conspiracy theories of QAnon to the incitement of white supremacists, social media are drowning users in hatred and falsehood. The right accuses the tech firms of censorship, including last week of a dubious article alleging corruption in the family of Joe Biden, the Democratic presidential nominee. And yet the question of what to do about social media is best seen through the same four stages as the case against Google: harm, dominance, remedies and delay. At stake is who controls the rules of public speech.

A tenth of Americans think social media are beneficial; almost two-thirds that they cause harm. Since February YouTube has identified over 200,000 "dangerous or misleading" videos on covid-19. Before the vote in 2016, 110m-130m adult Americans saw fake news. In Myanmar Facebook has been used to incite genocidal attacks against the Rohingyas, a Muslim minority. Last week Samuel Paty, a teacher in France who used cartoons of the Prophet Muhammad to talk about free speech, was murdered after a social-media campaign against him. The killer tweeted an image of Mr Paty's severed head, lying in the street.

The tech firms' shifting attempts to sterilise this cesspool mean that a handful of unelected executives are setting the boundaries of free speech (see [Briefing at the end of the Longer version of the selection on CdP](#)). True, radio and TV share the responsibility for misinformation and Republican claims of bias are unproven—right-wing sources often top lists of the most popular items on Facebook and Twitter. But pressure is growing on the tech firms to restrict ever more material. In America the right fears that, urged on by a Democratic White House, Congress and their own employees, the firms' bosses will follow left-leaning definitions of what is acceptable. Contrast that with the First Amendment's broad licence to cause offence.

Elsewhere, governments have also used social media companies to go beyond the law, often without public debate. In London the Metropolitan Police requests that they take down legal, but troubling, posts. In June France's Constitutional Council struck down a deal between the government and the tech companies because it curbed free speech—an initiative that is sure to be revisited after Mr Paty's murder. Citing Western precedents, more authoritarian governments in countries such as Singapore expect the tech firms to restrict “fake news”—potentially including irksome criticism from opponents.

This might not matter were the networks less dominant. If people could switch as easily as they change breakfast cereal, they could avoid rules they dislike. But switching is like giving up your mobile-phone number: it cuts you off from your friends. Social networks have also become so central to distributing news and opinion that they are, says Mark Zuckerberg, Facebook's founder, a “town square”. If you want to be part of the conversation you have no choice but to be there, soapbox in hand.

This hold over users has one further dismal implication for truth and decency. In order to sell more ads, the tech companies' algorithms send you news and posts that they think will grab your attention. Political cynics, con artists and extremists take advantage of this bias towards virality to spread lies and hatred. Bots and deep fakes, realistic posts of public figures doing or saying things that never happened, make their job cheaper and easier. They are rapidly becoming more sophisticated.

The purest remedy for this would be to change the tech firms' business model and introduce more competition.

That is already working well in other areas of tech, like the cloud. One idea is for people to own their data individually or collectively. If users could port their data to another network, the tech firms would have to compete to provide a good service.

The obstacles to this are immense. The tech firms' value would tumble by hundreds of billions of dollars. It is not clear you own the data about your online connections. You could not migrate to a new network without losing the friends who stayed behind unless the platforms were interoperable, as mobile-phone networks are. Perhaps the authorities could impose less sweeping remedies, such as giving users the right to choose feeds set by a neutral rule, not an attention-grabbing algorithm.

The keys to the hype house

Such ideas cannot be implemented quickly, but societies need solutions today. Inevitably, governments will want to set the basic rules at the national level, just as they do for speech. They should define a framework covering obscenity, incitement and defamation and leave judgments about individual posts to others. **International human-rights law is a good starting-point, because it leans towards free speech and requires restrictions to be relevant and proportionate, but allows local carve-outs.**

Social-media firms should take those standards as their basis. If they want to go further, attaching warnings to or limiting content that is legal, the lodestars should be predictability and transparency. As guardians of the town square, they ought to open their processes to scrutiny and particular decisions to appeal. Ad hoc rule changes by top executives, as with the recent Biden decision, are wrong because they seem arbitrary and political. Hard cases, like kicking opponents of Bashar al-Assad in Syria off a platform for mentioning terrorists, should be open to review by representative non-statutory boards with more power than the one Facebook has created. Independent researchers need much freer access to anonymised data so that they can see how platforms work and recommend reform. Such rule-making should be open to scrutiny. In America politicians can use removing the protection from prosecution granted by Section 230 of the Communications Decency Act as a lever to bring about change.

This will be messy, especially in politics. When societies are divided and the boundary between private and political speech is blurred, decisions to intervene are certain to cause controversy. The tech firms may want to flag abuses, including in post-election presidential tweets, but they should resist getting dragged into every debate. Short of incitement to violence, they should not block political speech. Politicians' flaws are better exposed by noisy argument than enforced silence.

By Geoffrey R. Stone, *Chicago Tribune*, Jan 12, 2021

A few days ago, shortly after the unprecedented invasion of the United States Capitol by a horde of unrestrained Trump supporters, Twitter announced that it was banning President Donald Trump from Twitter. After more than 36,000 tweets in four years, with 88 million followers, Trump was tossed, perhaps permanently, from his favorite social media megaphone.

Twitter explained that this was warranted because of Trump’s irresponsible use of Twitter to lie about the outcome of the 2020 presidential election and to incite the deadly riot that disrupted the critical deliberations of Congress, which was in the process of counting the electoral votes that would confirm Trump’s personally mortifying defeat. At approximately the same time and for the same reasons, Trump was also banned at least temporarily from Facebook.

Trump’s supporters immediately raged that these actions by Twitter and Facebook violated the First Amendment. Is this assertion correct? The answer is simple: No.

Like all of the guarantees of the Bill of Rights, the First Amendment’s fundamental guarantee of “freedom of speech, and of the press” limits *only* the actions of the government (federal, state and local), *not* the actions of private individuals, organizations or businesses. Just as a private enterprise cannot violate your constitutional right not to be subjected to “unreasonable searches and seizures,” just as a private enterprise cannot violate your constitutional right to the “free exercise of religion,” and just as a private enterprise cannot violate your constitutional right to terminate an unwanted pregnancy, so too is it legally impossible for Twitter or Facebook to violate your — or Trump’s — constitutional right to “freedom of speech, or of the press.”

A forgotten case

Having said this, I should note that in at least one case, decided some 75 years ago, the Supreme Court held in *Marsh v. Alabama* that a privately-owned company town could not constitutionally prohibit Jehovah’s Witnesses from distributing literature on the town’s sidewalks. In a once-in-history decision, the Court explained that “the more an owner, for his advantage, opens up property for use by the public in

general, the more do his rights become circumscribed by the ... constitutional rights of those who use it.”



Twitter Safety
@TwitterSafety

After close review of recent Tweets from the @realDonaldTrump account and the context around them we have permanently suspended the account due to the risk of further incitement of violence.

Twitter said on Jan. 8, 2021, that it had permanently suspended President Donald Trump from its service "due to the risk of further incitement of violence." (Twitter)

That decision, however, stands alone, and it has never been followed. If it had remained good law, it would be interesting to see how that principle would apply to Twitter and Facebook. But the decision in *Marsh* was a one-time anomaly. It is not the law, and for all practical purposes it has not been the law for three-quarters of a century.

But that leaves our nation with a puzzle. If, let us suppose, the United States government could not constitutionally pass a law banning Trump from using Twitter or Facebook, why should Twitter and Facebook be allowed to do so? Even if they are private enterprises and not the government, they have enormous power and influence over public discourse in our nation. Should they have unlimited authority to decide for themselves who can and cannot share their views with other Americans on these extraordinarily powerful means of communication? Can they — should they — be trusted to have such authority to determine the bounds of public discourse in our democracy?

Radio as a model

Our nation faced a somewhat similar question with the advent of radio in the 1920s. At that time, with only a small number of frequencies available in any location, the fear was that a small number of wealthy

individuals could buy up all the frequencies in a city such as Chicago and completely dominate this powerful new means of communication and then distort and corrupt our democracy.

With that concern in mind, Congress enacted the Communications Act of 1934, which established the Federal Communications Commission and granted it broad power to regulate the broadcast spectrum. The FCC then adopted the Fairness Doctrine, which imposed on radio and television broadcasters the requirement that discussion of public issues be presented in a fair and balanced manner.

In 1960, the Supreme Court upheld the constitutionality of the Fairness Doctrine, noting that “there is no sanctuary in the First Amendment for unlimited private censorship operating in a medium not open to all.” The Fairness Doctrine remained in effect until repealed by the Reagan administration in 1987.

The question we face going forward is whether social media should be subject to similar government regulations. Should Facebook, Twitter and other major social media platforms be required by law to operate

in a fair and balanced manner? Should they be required by law to function in a manner consistent with the First Amendment? Should they be required by law to ensure that individuals are exposed to both sides of a question in a responsible and evenhanded manner? Should they be constrained from taking down material that is protected by the First Amendment? This is a fundamental question our nation must wrestle with in the future.

If we move in that direction, what would that tell us about the decision to remove Trump from Twitter and Facebook? Would it be constitutional for the *government* to remove Trump from social media because of his past and recent posts? Were his consistently and dangerously false posts about the outcome of the election and about the protest/riots at the U.S. Capitol protected by the First Amendment, or did he go beyond the bounds of the First Amendment by intentionally inciting a likely and imminent seditious riot? That would be the critical question.

Geoffrey R. Stone is the Edward H. Levi Distinguished Professor of Law at the University of Chicago.

TEXT 5 -Donald Trump et les médias sociaux : « La question de la régulation privée du débat public ne date pas d’hier »

Entretien avec Romain Badouard Propos recueillis par *Xavier Eutrope*, La Revue des médias, INA, 14 janvier 2021

Depuis le 6 janvier, de nombreux réseaux sociaux ont banni temporairement ou définitivement Donald Trump, montrant le pouvoir déterminant des grandes plateformes numériques sur la liberté d'expression et le fonctionnement du débat public en Occident. Quelles sont les conditions d'une meilleure régulation de ces acteurs et de l'espace public numérique ?

Le mercredi 6 janvier, une foule compacte de militants pro-Trump radicalisés partent à l'assaut du Capitole, après y avoir été incités par Donald Trump. Le bilan à la fin de la journée est dramatique : cinq morts, dont un agent de la police du Capitole tué à coups d'extincteur et une partisane de Donald Trump mortellement blessée par un coup de feu. Dans les jours qui suivent, Le président américain en exercice voit ses comptes personnel suspendus (Facebook et YouTube) ou supprimés (Twitter). Raison invoquée : le président aurait enfreint leurs politiques de lutte contre l'incitation à la

violence. Ce « bannissement », inédit pour un chef d'État, provoque de nombreux débats sur la régulation des médias sociaux et la modération des propos qui y sont tenus. Entretien avec le chercheur Romain Badouard, maître de conférences en sciences de l'information et de la communication à l'université Paris-II et auteur des *Nouvelles Lois du Web. Modération et censure* (Seuil, 2020).

Comment analysez-vous la décision de Twitter, Facebook, Youtube et d'autres, de suspendre voire bannir Donald Trump de leurs plateformes ?

Romain Badouard : Cette affaire montre le pouvoir déterminant des grandes plateformes numériques sur la liberté d'expression ainsi que le fonctionnement du débat public et de la démocratie en Occident. Ces plateformes sont des entreprises privées, fournissant des services privés, et on peut considérer qu'à ce titre elles ont le droit de décider de la manière dont on utilise les technologies

qu'elles mettent à disposition, donc de modérer ou de bannir des utilisateurs, quels qu'ils soient. Mais ces espaces privés sont devenus les principales arènes du débat public dans les démocraties occidentales. On ne peut donc pas considérer que ces plateformes ne sont que des espaces privés. Elles doivent aussi faire l'objet d'une régulation de la part des pouvoirs publics.

De nombreuses personnes ont fait part de leur étonnement voire sidération à la suite du bannissement de Donald Trump de ces plateformes. Ce n'est cependant pas la première fois qu'elles prennent de telles décisions, unilatérales et discutables. Vous citez dans votre livre le cas de pages Facebook plutôt classées à gauche qui ont vu leur visibilité s'évanouir sans raison apparente.

Un palier a été franchi, et c'est ce qui attire autant l'attention. C'est le président des États-Unis, élu démocratiquement, qui se fait censurer et bannir par des modérateurs et des plateformes privés. Mais effectivement, la question de la régulation privée du débat public et du pouvoir de censure de ces plateformes ne date pas d'hier. Je me réjouis tout de même que nous en parlions autant. Ces controverses nous poussent à nous demander collectivement ce que pourrait être une régulation réellement démocratique de ces plateformes.

En lisant votre livre, on peut avoir le sentiment que ces services ont été fondamentalement mal construits, que leurs fonctionnalités n'ont pas forcément été pensées pour limiter les abus et permettre, par exemple aux personnes victimes de harcèlement, de se protéger. Est-il possible de les améliorer ?

Ces plateformes n'ont pas été créées pour devenir les arènes du débat public. Elles étaient initialement des réseaux sociaux qui devaient permettre à des personnes qui se connaissent de se réunir entre elles. Avec l'accroissement de leur popularité, elles sont devenues des espaces de débats très importants et ont, à ce titre, de nouvelles responsabilités. Les technologies ne sont jamais immuables et elles peuvent être déconstruites pour être reconstruites à nouveau. Ces plateformes peuvent être poussées à modifier leur modèle économique et/ou technologique. Il faudrait leur imposer des choix de design, notamment dans leur conception même. Instagram a par exemple expérimenté le fait de ne plus comptabiliser les

likes sur les photos. YouTube a décidé de moins recommander de contenus complotistes sur sa plateforme. Il y a beaucoup de leviers de régulation technologiques à notre disposition. Ce qui manque, c'est une volonté politique pour imposer ces choix, même si elle est aujourd'hui en train d'émerger.

Comment expliquez-vous que cette volonté politique a mis tant de temps à émerger ?

Les personnalités politiques n'ont peut-être pas perçu les dangers démocratiques derrière ces grandes plateformes, et ont d'abord vu ces réseaux comme des outils de modernité. On peut poser l'hypothèse qu'il fallait y être et ne surtout pas les critiquer, de peur de paraître vieux jeu. Ces réseaux étaient aussi vus comme des leviers de croissance économique, que l'on n'a pas voulu entraver alors qu'elle était promise par les grandes entreprises du numérique. Nous sommes entrés dans une nouvelle séquence depuis le milieu des années 2010 : ces grandes plateformes présentent une somme de dangers démocratiques potentiels, que ce soit la surveillance généralisée, de nouvelles formes de propagande, la réutilisation des données personnelles, la violence expressive dans le débat. C'est un peu le revers de la médaille. On arrive aujourd'hui à un moment historique : celui de la régulation.

Y a-t-il une volonté des pouvoirs publics de « déjudiciariser » ou extra-judiciariser les décisions concernant ce qui se dit et ce qui se fait sur les plateformes de médias sociaux ?

Cette question de l'extra-judiciarisation des décisions est un gros problème. Il est notamment dénoncé par les organisations de défense des libertés sur Internet, qui y voient des formes de privatisation de la censure. Vous avez des États qui font pression sur des acteurs privés pour censurer davantage ce qui peut se dire sur leurs plateformes, faisant disparaître, de fait, le juge du paysage. C'est notamment ce que prévoyait la loi Avia. Au passage, la loi sur les manipulations de l'information [parfois appelée « loi fake news » ou « loi infox », NDLR] en France ne s'inscrit pas dans cette tendance car elle crée une nouvelle procédure judiciaire.

Dans un monde idéal, un parquet numérique serait notifié de toutes les décisions des plateformes, et capable à la fois d'engager des procédures contre les internautes qui ont tenu des propos illégaux et de sanctionner les plateformes qui se livrent à de la censure abusive. Dans la pratique, cela paraît difficilement envisageable étant donné les moyens

dont dispose la justice aujourd'hui et le volume de contenus qui sont publiés sur les réseaux sociaux. Il faut donc plutôt réfléchir à la façon dont les plateformes vont mieux collaborer avec la justice pour permettre, grâce à la communication des données, l'identification de personnes qui se livrent à du cyberharcèlement, à de la propagande jihadiste, des menaces de mort, et ainsi de suite.

Il y a eu des procès pour des propos violents tenus et diffusés sur les réseaux sociaux. La justice est-elle efficace pour lutter contre ce genre de menaces ?

Les procès ont une efficacité symbolique qui ne doit pas être négligée. Il y a des procès médiatiques, je pense notamment à celui de Nadia Daam contre ses cyber harceleurs. Certes, c'est une goutte d'eau dans l'océan, mais ça a quand même le mérite de montrer que, non, on ne peut pas tout dire même lorsque l'on se croit protégé par un écran d'ordinateur. On ne peut pas menacer quelqu'un de mort, c'est interdit par la loi et cela vaut aussi pour Internet. Si cela peut permettre une prise de conscience de certains internautes, qui cherchent à faire des « mauvaises blagues », ce n'est peut-être pas plus mal. Ce que montrent aussi ces affaires, c'est qu'il est possible d'identifier ceux qui profèrent des menaces de mort ou harcèlent d'autres utilisateurs sur les réseaux sociaux et que même si ces personnes sont anonymes, elles peuvent être identifiées grâce à leur adresse IP, qui peut être transmise à la justice. Ce dont nous avons besoin, et c'est ce à quoi devrait s'atteler un parquet numérique, c'est que les plateformes collaborent de manière plus efficace avec le pouvoir judiciaire. Ce n'était, jusqu'à aujourd'hui, pas encore le cas. La lutte antiterroriste a montré que malgré tout, il était possible de le faire.

Comment faire pour que les plateformes répondent d'avantage aux exigences démocratiques plutôt qu'à des sanctions économiques, en sachant qu'il n'y a pas forcément d'alignement entre ces deux intérêts ? (Voir notre [interview de Yochai Benkler](#))

Les pouvoirs publics doivent fixer des normes minimales de fonctionnement de ces réseaux dans les démocraties. Mais la régulation ne peut pas se limiter uniquement au rôle des États : une régulation réellement démocratique des plateformes passe aussi par l'engagement des citoyens, celui des entreprises et par la régulation du marché publicitaire. On se rend compte que

beaucoup d'entreprises qui financent les espaces publicitaires sur ces plateformes ont un moyen de pression très important sur leur fonctionnement. Lorsqu'elles refusent de diffuser de la publicité dans des environnements dans lesquels elles jugent qu'il y a trop de désinformation ou trop de violence, des actions sont prises. On peut regretter qu'il faille en arriver à des menaces économiques pour faire bouger les choses, mais on peut aussi considérer que ces entreprises ont leur rôle à jouer dans la régulation des plateformes et de l'espace public numérique, notamment par le biais du fonctionnement du marché publicitaire, qui pourrait lui aussi être largement assaini sur internet.

On observe parfois un manque de transparence et de clarté à ce sujet. Vous évoquez un épisode particulièrement marquant dans votre livre. Des sociétés ont fait part à Google de leur mécontentement de voir certaines de leurs publicités apparaître sur des pages faisant l'apologie du terrorisme, les finançant de fait. Google a alors changé son fonctionnement, permettant aux annonceurs de choisir sur quels genres de pages leurs publicités pourraient apparaître... possibilité totalement transformée quelques mois après, sans que personne n'en sache rien, opérant de fait un retour en arrière.

La question de la régulation des plateformes par la transparence est essentielle, mais elle présente aussi un certain nombre d'écueils. On voit au sein des États européens, en ce moment particulièrement, des injonctions faites aux plateformes d'être plus transparentes en publiant des rapports présentant le fonctionnement de leurs dispositifs de modération, des registres des publicités politiques, etc. Ce qui est plutôt positif. En France, c'est le CSA qui doit contrôler ces rapports. C'est une bonne chose car cela permet d'en apprendre davantage sur la manière dont les plateformes modèrent les propos qui sont tenus. Mais réguler par la transparence pose un problème majeur : le régulateur devient dépendant des données qui lui sont transmises par les plateformes. On ne sait d'elles et de leurs actions que ce qu'elles veulent bien nous dire. Les plateformes fournissent des données, qu'il est impossible d'authentifier ou de certifier. Le nœud du problème consiste donc à savoir comment accéder de manière indépendante aux données des plateformes. La Commission européenne par exemple, avec le Digital Service Act qui a été présenté en décembre 2020,

préconise pour les très grandes plateformes l'obligation de réaliser des audits indépendants une fois par an, durant lesquels des agences spécialisées pourront avoir accès aux données des plateformes pour pouvoir les certifier. Sans accès certifié à ces données, la régulation par la transparence ne peut pas fonctionner.

Le caractère international des grandes plateformes numériques complique-t-il la situation ?

Pendant longtemps, ce caractère international d'Internet, avec des serveurs qui se trouvent dans un pays, des services qui sont proposés dans un autre, a compliqué beaucoup de choses, notamment la coopération judiciaire qui visait à limiter les discours illégaux dans certains pays. Aujourd'hui, paradoxalement, la centralisation du débat autour de grandes plateformes facilite les choses. Certes ces plateformes dépendent avant tout du droit américain, et pour la plupart du droit californien, mais on voit que les États, notamment en Europe, peuvent leur imposer des normes. Et les plateformes se mettent en conformité avec des dispositifs de modération à deux étages. Le premier étage, ce sont les standards de publication qui s'appliquent à l'ensemble des internautes dans le monde entier, le deuxième, ce sont les réglementations et restrictions nationales. Par exemple, l'adoption de la **loi NetzDG** en Allemagne ou de la **Loi sur les manipulations de l'information** en France, font que les versions des différents services ne sont pas exactement identiques, non seulement entre ces deux pays, mais aussi avec leurs voisins. Cela se manifeste différemment en fonction des plateformes. Sur Facebook, ce sont les types de contenus interdits qui diffèrent, sur YouTube, c'est la manière de signalement des contenus. Et sur Twitter, c'est l'utilisation de certains hashtags qui va être interdite ou autorisée en fonction des pays. Les plateformes sont dans une logique de mise en conformité avec les règlements nationaux, et c'est une bonne chose parce que cela montre que le pouvoir politique et le droit ont encore leur mot à dire dans la régulation de ces espaces.

Vous avez évoqué la régulation démocratique des plateformes, pourriez-vous nous parler de ce qui est en préparation au Royaume-Uni, où doit voir le jour une structure de surveillance de la régulation ?

L'une des grandes questions de la régulation des plateformes est de savoir qui doit s'occuper de cette

supervision de la régulation. En France il a été décidé que c'était le CSA. En Allemagne, c'est le ministère de la Justice qui mandate des organisations qui s'en occupent. **Au Royaume-Uni, la nouvelle législation vise à créer une autorité indépendante qui aurait pour mission la régulation des réseaux sociaux.** Elle serait créée à partir d'un impôt spécifique, levé sur le monde des entreprises de l'économie numérique. Il y a différentes façons de faire, tout l'enjeu étant de pouvoir accéder aux données des plateformes et de savoir comment nous serons capables de certifier que les informations qu'on contrôle sont bien authentiques.

En lisant votre livre, on peut se demander s'il ne serait pas plus simple de démonter toutes les plateformes et de les reconstruire de zéro sur la base d'initiatives publiques transnationales. Serait-ce souhaitable ou même réalisable ?

On peut imaginer que des grandes plateformes publiques et transnationales soient créées pour prendre la succession de ces grandes plateformes privées. Mais je me fais assez peu d'illusions sur le fait qu'elles ne rencontreraient pas de succès. Les plateformes comme YouTube, Facebook ou Twitter bénéficient de l'effet de réseau, un principe économique qui veut que l'on a intérêt à utiliser un service parce qu'il est utilisé par beaucoup de monde. Si j'utilise Facebook, c'est parce que tous mes amis y sont. Si je vais sur YouTube, c'est parce qu'il y a beaucoup de contenus disponibles. Créer des plateformes publiques respectueuses des données et qui ne seraient pas dans une démarche de ciblage publicitaire est tout à fait envisageable. On peut cependant penser que les internautes préféreront rester sur les plateformes existantes. A mon avis, la menace pour ces plateformes de médias sociaux vient plutôt du marché et de l'émergence d'autres services qui vont les concurrencer. On le voit aujourd'hui d'ailleurs avec Facebook, qui est beaucoup moins utilisé par les jeunes générations, qui lui préfèrent Snapchat ou TikTok. Ce que nous apprend la courte histoire du Web, c'est que les grandes entreprises qui sont en situation de quasi-monopole peuvent très bien tomber dans l'oubli en quelques années.

Est-il possible aujourd'hui de parvenir à une modération des contenus efficace, transparente et qui respecte les utilisateurs ?

C'est envisageable, mais la massification des publications représente un gros problème. Lorsque l'on voit, par minute, le nombre d'heures

de vidéos postées sur YouTube, le nombre de contenus publiés sur Facebook ou de messages partagés sur Twitter, on a du mal à imaginer comment une modération uniquement humaine pourrait en venir à bout. C'est pour cela que les plateformes ont opté pour des formes de modération automatisées, avec des algorithmes qui scannent les contenus à la recherche de publications problématiques, d'usage de mots ou d'images interdits. **Il me semble difficile d'aller contre ce phénomène d'automatisation.** Ces algorithmes créent aussi beaucoup de faux positifs et génèrent de la censure abusive, ce qui est un autre problème.

Face à ça, il faut garantir un certain nombre de droits aux usagers, notamment une transparence de la procédure de modération. Si, par exemple, nous sommes l'objet d'une procédure, il faut que nous soyons mis au courant, informé des décisions qui ont été prises et les critères sur lesquelles elles se fondent. Il faut aussi pouvoir bénéficier d'un droit d'appel qui nous permette de demander ou d'exiger que leurs contenus soient à nouveau évalués. Facebook, par exemple, a mis en place un droit d'appel depuis 2018. Les chiffres fournis par la plateforme elle-même montrent qu'il y a quand même une part non négligeable de contenus supprimés qui sont restaurés après appel, de l'ordre de 12 %. Il y a aussi des besoins à ce niveau-là.

Vous évoquez dans votre essai la formation de groupes de contre-discours sur les réseaux sociaux, qui veulent occuper le terrain et contrer les paroles haineuses en apportant des arguments ou en désamorçant l'agressivité des propos. De telles initiatives doivent-elles être financées publiquement ?

La régulation vraiment démocratique passe aussi par l'engagement des citoyens et leur participation. C'est une partie de la solution, mais ce n'est pas une solution miracle. Le contre-discours permet, dans bien des cas, de pacifier les débats ou d'alerter la majorité silencieuse, qui regarde les débats sans y participer, du caractère falsifié d'une information. C'est important, d'autant plus que les internautes sont au plus près du contenu. Pour les plateformes, ils correspondent aussi à une sorte de force de modération distribuée, qu'il est important de mobiliser. Les plateformes essaient par ailleurs de récupérer les initiatives de contre-discours, en cherchant notamment à financer des associations. Du côté des pouvoirs publics, on a récemment pu voir Marlène Schiappa appuyer la création de

cellules de riposte, qui porteraient des contre-discours républicains.

Beaucoup de gens ont fait part de leur circonspection face à cette proposition de la Ministre déléguée chargée de la citoyenneté...

Ce qui fait la force de ces contre-discours, à mon avis, c'est aussi leur caractère spontané. On peut imaginer que cela fonctionne un petit peu mieux lorsqu'il s'agit de conversations entre internautes que lorsque c'est labelisé par les plateformes ou par les pouvoirs publics. Lorsque Marlène Schiappa évoque cette initiative, on imagine des gens dans un ministère en train de répondre aux internautes, ce qui ne peut pas être efficace. Finalement, l'idée est moins de convaincre les internautes qui ne sont pas d'accord que d'occuper l'espace du débat et de montrer aussi que d'autres discours sont possibles.

En conclusion de votre essai, vous évoquez trois principes qui permettraient de protéger la liberté d'expression et la garantir. Pourriez-vous nous en dire un peu plus ?

Face à ce pouvoir des grandes entreprises privées sur la régulation du débat public, nous pouvons leur imposer des normes, mais nous pouvons aussi chercher à réfléchir à la manière dont nous pourrions doter les citoyens de nouveaux droits. L'un des principes que je présente en conclusion est celui de la transparence, c'est-à-dire comment imposer aux plateformes la publication régulière d'un certain nombre d'information liées à leurs activités de modération. Il y a ensuite le principe d'appel, qui veut que les internautes qui pensent avoir été victimes de censure abusive puissent exiger un réexamen de leur publication. Le dernier principe, c'est celui qui garantit le fonctionnement des deux autres : l'accès aux données. Des agences indépendantes ou des pouvoirs publics pourraient avoir un pouvoir d'audit qui permettrait de certifier les données transmises par les plateformes. C'est vraiment le nœud du problème. En France, les plateformes de médias sociaux transmettent des rapports au CSA ou publient des rapports sur des plateformes dédiées, mais nous savons très peu de choses sur l'authenticité de ces données.

Êtes-vous confiant ? Pensez-vous que les choses peuvent changer dans un futur proche ?

Je suis, bizarrement, plutôt confiant. Nous nous trouvons à un moment particulier qui est propice à la régulation de ces grandes plateformes. D'abord, parce que les pouvoirs publics, aux États-Unis ou en Europe, votent de nouvelles lois ou engagent des

procédures judiciaires qui poussent les plateformes à réformer leurs dispositifs de modération, ce qui fait bouger les choses. Et les médias, ainsi que les citoyennes et les citoyens, s'intéressent au sujet. Le fait que toutes et tous se saisissent du dossier crée un environnement propice. Jusqu'où cela ira-t-il,

arriverons-nous à des formes véritablement démocratiques de régulation ? C'est encore la grande question, mais on peut déjà se satisfaire que les choses bougent enfin, après de nombreuses années d'inertie.

Existing or pending legislation

EU Digital Markets Act and Digital Services Act explained

<https://www.europarl.europa.eu/news/en/headlines/society/20211209STO19124/eu-digital-markets-act-and-digital-services-act-explained>
03-05-2023

Parliament adopted two major pieces of legislation that will change the digital landscape: find out about the Digital Markets Act and Digital Services Act.

The landmark digital rules, adopted on 5 July 2022, will create a safer, fairer and more transparent online environment.

The power of digital platforms

Over the last two decades, digital platforms have become an integral part of our lives - it's hard to imagine doing anything online without Amazon, Google or Facebook.

While the benefits of this transformation are evident, the dominant position gained by some of these platforms gives them significant advantages over competitors, but also undue influence over democracy, fundamental rights, societies and the economy. They often determine future innovations or consumer choice and serve as so-called gatekeepers between businesses and internet users.

To address this imbalance, the EU is upgrading the current rules governing digital services by introducing the Digital Markets Act (DMA) and the Digital Services Act (DSA), which will create a single set of rules applicable across the EU.

Regulating big tech practices: Digital Markets Act

The purpose of the Digital Markets Act is to ensure a level playing field for all digital companies, regardless of their size. The regulation will lay down clear rules for big platforms - a list of "dos" and "don'ts" - which aim to stop them from imposing unfair conditions on businesses and consumers. Such practices include ranking services and products offered by the gatekeeper itself higher than similar services or products offered by third parties on the gatekeeper's platform or not giving users the possibility of uninstalling any preinstalled software or app.

Interoperability between messaging platforms will improve - users of small or big platforms will be able to exchange messages, send files or make video calls across messaging apps.

The rules should boost innovation, growth and competitiveness and will help smaller companies and start-ups compete with very large players.

The Digital Markets Act will also set out the criteria for identifying large online platforms as gatekeepers and will give the European Commission the power to carry out market investigations, allowing for updating the obligations for gatekeepers when necessary and sanctioning bad behaviour.

Safer digital space: Digital Services Act

The Digital Services Act will give people more control over what they see online: users will have better information over why specific content is recommended to them and will be able to choose an option that does not include profiling. Targeted advertising will be banned for minors and the use of sensitive data, such as sexual orientation, religion or ethnicity, won't be allowed.

The new rules will also help protect users from harmful and illegal content. They will significantly improve the removal of illegal content, making sure it is done as fast as possible. It will also help tackle harmful content, which, like political

or health-related disinformation, doesn't have to be illegal, and introduce better rules for the protection of freedom of speech.

The Digital Services Act will also contain rules making sure that products sold online are safe and follow the highest standards set in the EU. Users will have better knowledge of the real sellers of products that they buy online.

Next steps

The Digital Services Act came into force on 16 November 2022 and will be directly applicable across the EU by 17 February 2024. However, very large platforms and very large online search engines will have to comply with their obligations under the Digital Services Act earlier - maximum four months after the European Commission designated them. The Commission designated the first set of very large platforms on 25 April 2023.

The Digital Markets Act entered into force on 1 November 2022 and its rules started to apply on 2 May 2023. The European Commission will designate gatekeepers by 6 September 2023 at the latest and they will then have a maximum of six months to comply with the new obligations under the Digital Markets Act, so by March 2024.

See also: <https://www.europarl.europa.eu/news/en/headlines/economy/20201022STO89919/meps-ideas-for-tackling-harmful-or-illegal-content-online>

In the UK – The Online Safety Bill

The **Online Safety Bill** is a proposed Act of the Parliament of the United Kingdom intended to improve internet safety, published as a draft on 12 May 2021.

The Bill would create a new duty of care for online platforms towards their users, requiring them to take action against both illegal and legal but harmful content. Platforms failing this duty would be liable to fines of up to £18 million or 10% of their annual turnover, whichever is higher. It would also empower **Ofcom*** to block access to particular websites. Additionally, the Bill would oblige large social media platforms not to remove, and to preserve access to, journalistic or "democratically important" content such as user comments on political parties and issues.

The bill has been heavily criticised for its proposals to restrain the publication of "lawful but harmful" speech, effectively creating a new form of censorship of otherwise legal speech. As a result, in November 2022, measures that were intended to force big technology platforms to take down "legal but harmful" materials were removed from the Online Safety Bill. Instead, tech platforms will be obliged to introduce systems that will allow the users to better filter out the harmful content they do not want to see

See here for more: <https://www.gov.uk/guidance/a-guide-to-the-online-safety-bill#a-guide-to-the-online-safety-bill>

* **OFCOM:** The **Office of Communications**, commonly known as **Ofcom**, is the government-approved regulatory and competition authority for the broadcasting, telecommunications and postal industries of the United Kingdom. Ofcom has wide-ranging powers across the television, radio, telecoms and postal sectors. It has a statutory duty to represent the interests of citizens and consumers by promoting competition and protecting the public from harmful or offensive material

In France

● ARCOM

L'**Autorité de régulation de la communication audiovisuelle et numérique (Arcom)** est l'autorité publique indépendante française qui résulte de la fusion le 1^{er} janvier 2022 du Conseil supérieur de l'audiovisuel (CSA) et de la Haute Autorité pour la diffusion des œuvres et la protection des droits sur internet (Hadopi)

<https://www.arcom.fr/>

<https://www.arcom.fr/nos-missions/regulation-des-plateformes-en-ligne-et-reseaux-sociaux>

● **Loi organique et loi ordinaire du 22 décembre 2018 relatives à la manipulation de l'information - www.vie-publique.fr**

La loi a été promulguée le 22 décembre 2018. Elle a été publiée au Journal officiel du 23 décembre 2018.

Les lois visent à lutter contre la manipulation de l'information à l'heure numérique et à endiguer la diffusion de fausses informations ("fake news") pendant les périodes de campagne électorale.

Elles créent une nouvelle voie de référé civil visant à faire cesser la diffusion de fausses informations durant les trois mois précédant un scrutin national. Quand il est saisi, le juge des référés doit apprécier, sous 48 heures, si ces fausses informations sont diffusées "de manière artificielle ou automatisée" et "massive".

Dans sa décision du 20 décembre 2018, le Conseil constitutionnel a précisé que le juge ne pouvait faire cesser la diffusion d'une information que si le caractère inexact ou trompeur de l'information était manifeste et que le risque d'altération de la sincérité du scrutin était également manifeste.

Les plates-formes numériques (Facebook, Twitter, etc.) sont soumises à des obligations de transparence lorsqu'elles diffusent des contenus contre rémunération. Celles qui dépassent un certain volume de connexions par jour doivent avoir un représentant légal en France et rendre publics leurs algorithmes.

Le **Conseil supérieur de l'audiovisuel (CSA)** peut aussi empêcher, suspendre ou interrompre la diffusion de services de télévision contrôlés par un État étranger ou sous l'influence de cet État, et portant atteinte aux intérêts fondamentaux de la nation.

In the US – No federal regulatory body per se

Here is an extract from <https://studentbriefs.law.gwu.edu/ilpb/2021/12/08/online-content-regulation-an-international-comparison/>

The United States, on the other hand, has a long tradition of permitting speech that is prohibited by other countries, including other liberal democracies. Under US First Amendment law, most speech, including hateful and violent speech, is protected, unless the speaker is using “fighting words” or inciting “imminent lawless action” that could lead directly to violence, as per the US Supreme Court case *Brandenburg v. Ohio* (1969). The American tradition of free speech has influenced some other jurisdictions, such as India, where the courts have traditionally held that in order to ban hateful and violent speech, the government must show a proximate and direct connection between the speech and the imminent violence. However, for the most part, the American approach toward speech goes against the grain of much of the rest of the world’s approach. Even in other liberal democracies, the dignity of the individual, or the security and interests of the state, are held to be more important than the absolute freedom of speech.

The American approach to speech has strongly influenced its domestic regulatory regime toward online harms, which can be classified as a broad immunity approach. Under **Section 230 of the Communications Decency Act (1996)**, often referred to as CDA 230, and subsequent case law, almost all online speech is legal in the United States. There are some specific legislative carve outs for sex trafficking, child pornography, and copyright violations, but otherwise, platforms cannot be held legally liable for the hateful content, defamatory speech, and breaches of privacy posed by individuals. Therefore, platforms where such speech may be posted—including Facebook, Twitter, and other sites—cannot be held liable for third-party posts.

This has been the subject of much domestic debate in recent years, with a bipartisan consensus emerging over the past few years that the United States Congress needs to amend or modify CDA 230 and the online harms regime. Members of both the Republican and Democratic parties want to reduce the scope of CDA 230 on the categories of “child sexual exploitation, content moderation operations...[and] content that courts determine to be illegal.”^[1] The two parties diverge, however, on how to change CDA 230 in other ways. Many on the right want to combat political censorship, while many on the left want to encourage platforms to prevent hosting content that “‘incite[s] or engage[s] in violence, intimidation, harassment, threats, or defamation’ against others based on various demographics, including race, religion and gender identity.”

However, at the present, individual companies may enforce their own community standards, which has ignited much domestic controversy over what can or cannot be posted on platforms used by hundreds of millions of Americans. While the standards used by Facebook and Twitter to remove hate speech are more similar to global norms, they are more restrictive than what is required by US law and First Amendment jurisprudence. Therefore, online social media platforms are increasingly subject to a battle between two norms: freedom of speech, and freedom from hate.

The American approach to online regulation of harmful content raises an interesting question: to what extent is government regulation of media content desirable, and would that regulate strengthen or weaken freedom of expression? The answers vary by jurisdiction.

In Germany

La loi Netzwerkdurchsetzungsgesetz, ou **NetzDG** est une loi allemande destinée à sanctionner les *fake news* et contenus haineux sur les réseaux sociaux. La loi a été votée le 30 juin 2017 par le Bundestag et est entrée en application le 1^{er} octobre 2017. La période de transition laissée aux réseaux sociaux pour s'adapter à la nouvelle loi a expiré le 1^{er} janvier 2018.

See also: https://www.lemonde.fr/pixels/article/2019/07/04/lutte-contre-la-haine-en-ligne-l-impossible-bilan-de-la-loi-allemande_5485218_4408996.html

Takeover by Elon Musk and the Free Speech debate

TEXT 6 - How Elon Musk Is Changing the Twitter Experience

Nearly six months after buying Twitter, Mr. Musk has made tweaks that have altered what people see on the platform and how they interact with it.

By Kate Conger, *The New York Times*, April 7, 2023

Kate Conger, a tech reporter in San Francisco, has reported on Twitter since 2018.

Elon Musk has declared he wants to transform Twitter into an all-inclusive app that people can use for payments, news and food orders.

“Buying Twitter is an accelerant to creating X, the everything app,” Mr. Musk posted in October, weeks before completing a \$44 billion acquisition of the social network. He later said Twitter could be like WeChat, the popular Chinese app that combines social media, instant messaging and payment services.

But nearly six months after Mr. Musk took over Twitter, his ambitions for the platform have remained mostly that — ambitions.

Although the billionaire has made dozens of tweaks to Twitter, they have largely been cosmetic. His changes have mostly affected the platform’s appearance, said Jane Manchun Wong, an independent software engineer who studies social apps. Those updates include adding more symbols and metrics displayed with tweets, but Twitter’s main elements — making it a place to quickly share news and discuss live events — haven’t altered.

Still, users’ experiences are changing. That’s because the kinds of tweets that they see are being affected by Mr. Musk’s behind-the-scenes adjustments. He has tinkered with the algorithm that decides which posts are most visible, thrown out content moderation rules that ban certain kinds of tweets and changed a verification process that confirms the identities of users.

The upshot is a Twitter that looks similar to the way it always has, but that is clunkier and less predictable in what tweets are surfaced and seen, users said. In some cases, that has caused confusion. Even Twitter’s employees have expressed frustration.

Last month, Andrea Conway, a designer at Twitter, posted about the design changes, saying: “We know you hate it. We hate it too. We’re working on making it suck less.” The modifications, she added, could eventually make Twitter “completely unusable.”

Mr. Musk did not respond to a request for comment.

So what looks different on Twitter now, and what are the changes underlying the tweaks?

The Newsfeed

The most notable difference is Twitter’s newsfeed, the stream of posts that people see when they open the app. Newsfeeds previously appeared as a single flow of posts, displaying tweets from only the accounts that a user followed. Mr. Musk has cleaved the newsfeed into two. Now when users open Twitter’s app, they see an algorithmically curated “For You” feed, which mimics a popular feature on TikTok, and a “Following” tab.

The “For You” newsfeed incorporates changes that Mr. Musk has made to Twitter’s recommendation algorithm, pulling in more tweets from people a user doesn’t follow and suggesting new topics and interests. That also means users might

see posts from all sorts of content creators whom they might not be interested in. At one point in February, the algorithm flooded users' feeds with tweets from Mr. Musk.

Here's how a user's "For You" newsfeed might look, with an example of a tweet from an account that the user doesn't follow and that the algorithm suggests. For users to see posts only from people they follow, they would have to switch to the "Following" tab.

Check Marks

Mr. Musk has also modified Twitter by adding a flood of color-coded check marks, which belie a deeper change to how the platform confirms the identities of organizations, governments, notable individuals and other official accounts.

Twitter previously offered white-and-blue check marks for users who were "verified," a kind of badge for those who had substantiated their identities and who were typically public figures, such as politicians and celebrities. The check marks were free.

Mr. Musk has begun charging users an \$8 monthly fee in exchange for a check mark, with the free check marks starting to disappear this month. He is essentially favoring payments from subscribers, departing from the idea that a check mark meant an account was notable.

Now yellow check marks indicate corporate accounts, while gray check marks denote the accounts of government officials. Companies can also add their logo to employees' accounts, verifying their employment. Individuals who pay get the blue-and-white check mark.

Those who paid for check marks would be boosted by Twitter's recommendation algorithm and be eligible to appear in people's "For You" newsfeeds, Mr. Musk said last month. That would prevent spam accounts from gaming the algorithm and rising to the top of the "For You" newsfeeds, he added.

Metrics, Metrics, Metrics

For most of Twitter's history, users could only like, retweet or reply to a post. The numbers of replies, likes and retweets then showed up at the end of a tweet.

Under Mr. Musk, every tweet now has more metrics attached. He has added a tally showing how many times a post has been viewed, saying the total number of views can demonstrate a message's popularity better than its total likes or retweets. Twitter has also added a tally for the number of times a tweet is bookmarked and saved.

That means every tweet now has the number of replies, likes, retweets, bookmarks and views appended to it. Here's an example of how tweets looked before, when there were fewer numbers, and after, with more metrics shown.

What do all of these moves add up to? Not necessarily the smoothest experience, some Twitter users and employees said.

"Twitter has leaned in to the 'crazy uncle' contingent," said Chris Messina, who is known as the inventor of the hashtag, adding that he now sees recommended tweets that don't align with his interests. "In terms of the product, overall I think the quality has really regressed."

TEXT 7 - We're going to need an alternative to Twitter

By Perry Bacon Jr., Columnist

The Washington Post, October 21, 2022

We need someone to create a new Twitter-like social media service. Starting yesterday.

The Post's report Thursday that Elon Musk is considering cutting 75 percent of Twitter's workforce once his purchase of the company is finalized was alarming enough. I'm not exactly sure how many employees Twitter needs, but it's hard to imagine that the site will improve or even maintain its current functionality with such a massive staffing reduction.

But even if the job cuts don't turn out to be that severe, a Twitter under Musk's control presents a huge looming problem for people who value the platform.

Why? Because Twitter has become a hugely important forum for challenging the rich and powerful. Twitter is a gathering place for news junkies, where anyone can post and any post can go viral, and its existence has shifted power in a fundamental way.

Billionaires, politicians and columnists at papers like this once enjoyed virtually one-sided communication — they put out their ideas, and few people had much of a public platform to disagree with them. With Twitter, you can disagree or sharply criticize a powerful person to an audience of millions. And if enough people see it your way, that feedback loop becomes a lever of accountability.

The powerful don't like this, and Musk, of course, is part of that powerful set. I would expect a Twitter

under Musk's control to change in ways that make it harder to blast those on high in ways that reach a mass audience.

That's not the only problem. The Tesla founder, whose personal politics seem to be anti-progressive-left more than anything else, is almost certain to end the Twitter bans on former president Donald Trump and other right-wing people who sometimes encourage violence and speak in racist and sexist ways on the site. I am a bit leery of these bans, because I would like Twitter to be a forum for all kinds of ideas and ideologues. But I suspect Musk will be simply looking to make Twitter more conservative to annoy liberals — not to help it carve out a sweet spot where broad ideas thrive but violence and bigotry are discouraged.

Finally, Musk just doesn't appear to get the value of the platform for many of those who use it. The thing I like most about Twitter is that it provides me a way to follow writers, academics, activists and others who aren't regularly published in major outlets, in

part because their views break from the White centrist establishment tone of most major U.S. news institutions.

Musk doesn't seem much interested in the thoughts of anyone besides other wealthy people, conservatives and anti-left-wing Democrats. At best, he won't be a good steward of the virtues of Twitter. At worst he will severely weaken them.

I need something like Twitter. I don't need an Elon Musk-infused anything. Watching this story over the past few months, my only question is whether we should wait and see how Musk runs Twitter, or if some tech expert or entrepreneur should just start a new service now. I tend toward the latter.

Perry Bacon Jr. is a Washington Post columnist.

See also

<https://www.theguardian.com/commentisfree/2023/jan/21/back-to-the-future-how-mastodon-is-restoring-the-lost-art-of-online-conversation>

TEXT 8

Elon Musk's Twitter is fast proving that free speech at all costs is a dangerous fantasy

[Nesrine Malik](#), *The Guardian*, Mon 28 Nov 2022

Free speech absolutists are like the cocky audience of a spectator sport – they think they could do better than the players, if they were just allowed a crack at it. To them, speech should be as free as possible, period. Nowhere is their oversimplification of the issue more evident than on social media, where abuse and disinformation have created a new frontier of regulation – and with it a cohort of disingenuous free speech warriors.

These absolutists are so unaccustomed to facing consequences for their actions that they have pushed the idea that a censoring “woke” orthodoxy now prevails, and is a threat to freedom of expression. Elon Musk is among them, but since his takeover of Twitter he is having to learn quickly that free speech is not simply about saying whatever you want, unchecked, but about negotiating complicated compromises.

Musk arrived at Twitter with an approach that I am sure he thinks is pretty straightforward. The site, he believes, has a leftwing bias that should be corrected by allowing suspended users back on to the platform. The accounts of Donald Trump, Kanye West and Jordan Peterson have been reinstated, along with nearly all those that were suspended for falling foul of old Twitter's rules on abuse and hate speech.

This means that Twitter is about to turn into a far more unpleasant and potentially dangerous experience. Little of this appears to have anything to do with a political strategy on Musk's part. Like Trump, Musk has become the tribune of fascists and racists by way of adolescent contrarianism, an insatiable need to flaunt his control and a radicalising inability to cope with being told he's wrong on the internet. For him, “free speech” seems merely a vehicle for his delusional plan to make Twitter into a fawning “digital town square” that he presides over.

But not even the richest man in the world can pull that sort of free speech arena off. Twitter isn't sustained by previously suspended users, but by the millions of people for whom the platform feels (most of the time) like a political and cultural slipstream. Twitter has an odd social media profile. It is both extremely influential and also often quite trivial, and the coexistence of the two is what makes the site viable. Twitter is a window into the soul of politicians and opinion-makers – its style of interactive rolling commentary works well in drawing them out to post their views or engage with others, revealing personalities and politics that otherwise would be suppressed or

closely edited. And it is the first resort of citizen journalists and those marshalling political protest. It also remains the only social media platform where people with little clout or profile can challenge elites directly.

But Twitter is also a solipsistic place, where even small users can become protagonists in spats that are then amplified both by the site's algorithms and a rightwing media that trawls it for telltale signs of "wokeness" or "cancel culture". For better or worse, it is Twitter's adjacency to current affairs and general political and cultural discourse that makes it, uniquely among platforms, feel *relevant*.

If you're not on Twitter, chances are that you have come across stories that started out or were precipitated there, whether it's a debate on trans rights that swirls around JK Rowling's tweets, or calls to organise street protests against dictators in the Arab world. For all these things to be possible on the same site, robust content moderation is necessary to ensure conversations don't descend into doxing (maliciously publishing someone's personal information) and hateful conduct, and that news and journalism is verifiable. In the absence of moderation, or at least the appearance of it, things fall apart pretty quickly. When a place is not fun or hospitable or truthful to users, it also becomes commercially pointless for advertisers. Since Musk took over, half of Twitter's top 100 advertisers are reported to have left the site. If things continue as they are, it is hard to see a future for the company.

The ultimate cause of that demise will be the failure of Musk to understand that for some speech to be free, other speech has to be limited. It is generally true that if a service is free then it is by definition exploitative of its users – if you are not paying for a product, the axiom goes, then you *are* the product. But in the case of social media, the regulation of your speech is the product. If a platform becomes too toxic, then it is useless for anyone except those who want an extremist ghetto of agitators.

In that sense, social media is very much like society in general. Political and legal authorities are in the business of content moderation, in order to make our shared space as stable and safe as possible for a majority of people. The public and other stakeholders, such as the press, businesses and social media companies themselves, are in constant negotiations with these authorities on what those limits should be – for instance, whether religious dress is protected speech, or what constitutes incitement to violence.

Old Twitter was far from perfect, and by its own admission its algorithms favoured rightwing accounts. But it was improving because of the drag that advertisers, regulators and users were putting on its algorithmic urge to encourage antagonistic activity. Twitter's chance of survival is dependent on whether Musk chooses to accept that, like freedom of speech, his power is not absolute.

Nesrine Malik is a Guardian columnist

TEXT 9 - « Il n'y a que les naïfs pour croire que la défense de la liberté d'expression est la seule motivation d'Elon Musk pour racheter Twitter »

Chronique Stéphane Lauer, éditorialiste au « Monde »

Le milliardaire américain met en avant sa volonté de laisser libre cours aux conversations les plus dérangeantes, sans modération, pour justifier l'acquisition du réseau social. Sans toutefois rappeler les intérêts stratégique, politique et économique qui motivent ce rachat.

Le Monde, 02 mai 2022

Elon Musk tente de nous convaincre qu'il travaille sans relâche à l'amélioration du bien-être de l'humanité. Son joyau, Tesla, est en voie de nous affranchir avec succès de la mobilité liée au pétrole. Space X nous promet la colonisation de Mars, si toutefois nous avons besoin d'une planète B. Sa nouvelle start-up, Neuralink, ouvre, quant à elle, de nouvelles perspectives transhumanistes, dont la première étape sera de redonner la parole et la mobilité aux personnes paralysées. « *Bigger than life* », comme disent les Américains.

Voilà maintenant que les ambitions prométhéennes d'Elon Musk franchissent une nouvelle étape. Avec le rachat de Twitter pour 44 milliards de dollars (41,7 milliards d'euros), il s'érige désormais en protecteur de la liberté d'expression et en garant de « *l'avenir de la civilisation* ». Amen.

Le génial entrepreneur met la main sur le site de microblogging comme il l'a pratiqué jusqu'à présent en tant qu'utilisateur, c'est-à-dire à coups de provocations, d'insultes ciblées, de manipulations en tout genre et d'un mélange de cynisme et de démagogie qui n'appartient qu'à lui. Le milliardaire est l'un des trolls les plus actifs du Web.

Mais ce qui pouvait distraire ou agacer prend une tout autre allure aujourd'hui car Twitter n'est pas une entreprise comme une autre. Agora virtuelle où débattent les principaux leaders d'opinion au milieu de millions d'internautes, le réseau social influence le climat politique du moment, qu'on le veuille ou non.

Troll, Elon Musk l'est lorsqu'il essaye de nous faire croire que le rachat de Twitter n'aurait d'autre but que de restaurer la liberté d'expression dans des sociétés occidentales rongées par la polarisation politique. L'initiative a rapidement trouvé un écho favorable auprès de tout un pan de la droite conservatrice américaine, notamment dans sa frange la plus extrémiste et complotiste. Ce courant s'estime victime d'une forme de censure qui ne ferait que protéger les idées progressistes et le politiquement correct.

Le « camp du bien », comme le nomme avec ironie le milliardaire libertarien, a démarré au quart de tour en alertant sur le retour en arrière que constituerait une moindre modération des contenus de Twitter. Où s'arrête la liberté d'expression ? Qui la contrôle ? Doit-elle être totale, comme le réclame M. Musk, ou bien faut-il la filtrer pour écarter les contenus haineux, le harcèlement et les fausses nouvelles ? Autant de questions fondamentales qui trouveront difficilement des réponses définitives dans un contexte de clivage exacerbé, largement façonné par les réseaux sociaux eux-mêmes.

Là se situe le nœud du problème. Le véritable pouvoir d'une plate-forme comme Twitter n'est pas tant sa capacité à faire le tri dans les opinions, que de leur donner une audience considérable. Autrement dit, les débats ne doivent pas seulement se focaliser sur le fait qu'un propos est acceptable ou non, mais surtout sur l'amplification algorithmique dont ils bénéficient.

Les réseaux sociaux disposent de tous les outils pour mesurer et limiter ces effets de viralité, mais rechignent à s'en servir parce que ces amplificateurs sont au cœur de leur modèle économique. Si Elon Musk veut laisser libre cours aux conversations les plus dérangeantes, c'est aussi parce que c'est bon pour le business. Il n'y a que les naïfs pour croire que la défense de la liberté d'expression est sa seule motivation.

Comme l'explique l'universitaire américaine Shoshana Zuboff, autrice de *L'Age du capitalisme de surveillance* (Zulma essais, 2020), de la même façon que le cerveau peut développer une addiction au sucre, l'esprit humain a une propension à être attiré par ce qui choque, indigné ou surprend. Dès lors, la stratégie pour Twitter consiste à sélectionner le contenu le plus toxique et le plus polémique pour maximiser le temps de connexion, les réactions et les partages entre utilisateurs afin d'augmenter non seulement les revenus publicitaires, mais surtout la collecte de données. Si Elon Musk était sincère, il renoncerait de lui-même à cette logique économique.

Pouvoir exorbitant

L'extraction à grande échelle de ces données, leur appropriation systématique et la sophistication de leur analyse donnent à Facebook ou Twitter un pouvoir exorbitant. Ces entreprises sont en capacité d'orienter les discussions, les opinions et même les comportements dans le réel, comme la participation à des manifestations ou la façon de voter, le tout sans contrôle et subrepticement.

« Le patron de Facebook est assis à son clavier céleste, et il peut décider, heure par heure, si les gens vont être plus ou moins en colère et si leur expression doit exister ou non », indique Shoshana Zuboff au [Washington Post](#). Avec Elon Musk, c'est risquer de mettre ce pouvoir « sous stéroïdes », affirme-t-elle.

Le milliardaire a-t-il un agenda politique ? Rien ne permet de l'affirmer. Ses motivations pourraient être bien plus prosaïques. « Mettre la main sur Twitter, c'est accéder à une gigantesque base de données conversationnelle, qui peut se révéler très précieuse pour développer l'intelligence artificielle dans ses autres business », avance Olivier Ertzscheid, spécialiste des nouvelles technologies à l'université de Nantes.

Le patron de Tesla se présente comme un « absolutiste de la liberté d'expression », mais c'est surtout le pouvoir qu'il s'arroge grâce à sa gigantesque fortune qui va devenir absolu. Une fois Twitter retiré de la Bourse, le milliardaire n'aura plus de comptes à rendre. Il promet plus de transparence sur les algorithmes, mais quelle valeur accorder à cet objectif avec une telle gouvernance ? Elon Musk va décider de ce qui est bien ou mal, alors qu'il n'est pas prêt à remettre en cause une logique économique des réseaux sociaux qui a déjà fait d'énormes dégâts.

Sa formule-choc résonne comme un splendide oxymore qui ferait presque oublier que l'une des plus solides démocraties au monde a frôlé le pire, le 6 janvier 2021, à Washington, lors de l'assaut du Capitole, à force de croire que n'importe qui pouvait dire n'importe quoi, par n'importe quel moyen.

Sarah Glozer, Senior Lecturer in Marketing & Society, University of Bath
Emily Jane Godwin, PhD Candidate in Cyber Security, University of Bath
Rita Mota, Assistant Professor, Department of Society, Politics and Sustainability, ESADE

For a man who made a fortune from electric cars, the Twitter takeover has turned into a fairly bumpy ride so far. Soon after buying the social media company for US\$44 billion (£38 billion), Elon Musk said he had “no choice” about laying off a large proportion of the company’s staff.

He has already faced a backlash over his move to charge Twitter users a monthly fee for their “blue tick” verified status. And those users should also be concerned about plans from the self-proclaimed “free speech absolutist” to reduce content moderation.

Moderation, the screening and blocking of unacceptable online content, has been in place for as long as the internet has existed. And after becoming an increasingly important and sophisticated feature against a rising tide of hate speech, misinformation and illegal content, it should not be undone lightly.

Anything which weakens filters, allowing more harmful content to reach our screens, could have serious implications for human rights, both online and offline.

For it is not just governments which are responsible for upholding human rights – businesses are too. And when different human rights clash, as they sometimes do, that clash needs to be managed responsibly.

Social media has proved itself to be an extremely powerful way for people around the world to assert their human right to freedom of expression – the freedom to seek, receive and impart all kinds of information and ideas.

But freedom of expression is not without limits. International human rights law prohibits propaganda for war, as well as advocacy of national, racial or religious hatred that constitutes incitement to discrimination, hostility or violence. It also allows for restrictions necessary to ensure that rights or reputations are respected.

So Twitter, in common with other online platforms, has a responsibility to respect freedom of expression. But equally, it has a responsibility not to allow freedom of expression to override other human rights completely.

After all, harmful online content is often used to restrict the freedom of expression of others. Sometimes, online threats spill over to the offline world and cause irreparable physical and emotional harm.

Any moves to remove content moderation therefore risk breaching corporate human rights obligations. Unlimited freedom of expression for some almost inevitably results in the restriction elsewhere of that exact same freedom. And the harm is unlikely even to stop there.

Musk claims that Twitter will now become a more democratic “town square”. But without content moderation, his privately owned version of a town square could become dysfunctional and dangerous. Twitter – again, like most other social media platforms – has long been linked to overt expressions of racism and misogyny, with a flood of racist tweets even surfacing after Musk closed his deal.

And while Musk reassures us that Twitter will not become a “hellscape”, it is important to remember that content moderation is not the same as censorship. In fact, moderation may facilitate genuine dialogue by cracking down on the spam and toxic talk which often disrupt communication on social media.

User friendly?

Moderation also offers reassurance. Without it, Twitter risks losing users who may leave for alternative platforms considered safer and a better ideological fit. Valuable advertisers are also quick to move away from online spaces they consider divisive and risky. General Motors was one of the first big brands to announce a temporary halt on paid advertising on Twitter after Musk took over.

Of course, we do not yet know exactly what Musk’s version of Twitter will eventually look like. But there have been suggestions that content moderation teams may be disbanded in favour of a “moderation council”. If it is similar to the “oversight board” at Meta (formerly Facebook), content decisions are set to be outsourced to an external party representing diverse views. But if Twitter has less internal control and accountability, harmful content may become a harder beast to tame.

Such abdication of responsibility risks breaching Twitter’s human rights obligations, and having a negative impact both on individuals affected by harmful content, and on the overall approach to human rights adopted by other online platforms. (...)

Content moderation is by no way a panacea and the claim that social media platforms are “arbiters of the

truth” is problematic for many reasons. We must also not forget the emotional and psychological toll of human content moderators having to view “the worst of humanity” to protect our screens. Yet, sanitisation of

social platforms is also not the answer. The internet is a better place when the most successful platforms engage in human rights-focused screening – for everyone’s benefit.

You can also listen to this program on NPR:

<https://www.npr.org/2022/10/08/1127689351/elon-musk-calls-himself-a-free-speech-absolutist-what-could-twitter-look-like-un>

TEXT 11 - Elon Musk's about-face on government censorship

Analysis by Cristiano Lima, with research by David DiMolfetta

The Washington Post, May 15, 2023

For months, Twitter chief **Elon Musk** has vowed that the platform under his watch would not “censor” on behalf of the U.S. government, as he claimed it had done in the past.

In November, he pledged that “Twitter will not censor accurate information about anything.” Later that month, he called for a “revolution against online censorship in America.” In December, he suggested U.S. officials engaged in “hidden state censorship in direct violation of the Constitution of the United States.” In April, he tweeted, “Censor not, lest ye be censored.”

The months-long crusade culminated in the release of the so-called “Twitter Files,” a trove of leaked internal documents that Musk and conservative social media critics seized on as evidence the U.S. government sought to suppress speech online.

Musk claimed the records showed that the “Government paid Twitter millions of dollars to censor info from the public” — a claim fact-checkers said was unsubstantiated.

But over the weekend, Musk and the company disclosed plans to do the very thing he has spent months decrying: censor speech at the behest of a national government.

Twitter late Friday said that in “response to legal process and to ensure Twitter remains available to the people of Turkey,” it had “taken action to restrict access to some content in Turkey today.”

The company announced the move just hours before a landmark and tightly contested election kicked off in Turkey on Sunday pitting longtime President **Recep Tayyip Erdogan** against challenger **Kemal Kilicdaroglu**.

As my colleague **Perry Stein** reported, critics have accused Erdogan of “cracking down on social media companies to stifle opposition voices as he tries to stay in power.”

Twitter did not publicly say what accounts would be impacted but that they had “informed” those users. Musk said Saturday they would disclose “what the government in Turkey sent us.”

The announcement sparked immediate backlash, with critics saying it contradicted Musk’s stated aim to lead the company as “a free speech absolutist.”

“Given Twitter’s total lack of transparency, it is hard to avoid the conclusion that Musk’s promises of free speech have again fallen away,” tweeted Rep. **Adam B. Schiff** (D-Calif.), the former chair of the House Intelligence Committee.

“The Turkish government asked Twitter to censor its opponents right before an election and @elonmusk complied — should generate some interesting Twitter Files reporting,” Bloomberg Opinion columnist and former Vox writer **Matthew Yglesias** tweeted Saturday.

Musk took umbrage at the comments, casting it as a difficult but necessary action. “Did your brain fall out of your head, Yglesias?” he tweeted back. “The choice is [sic] have Twitter throttled in its entirety or limit access to some tweets. Which one do you want?”

Twitter’s decision in Turkey, however, is part of a larger trend of the company buckling to government takedown demands more often since Musk took over.

According to a report last month by tech news publication Rest of World, Twitter under Musk “has complied with hundreds more government orders for censorship or surveillance — especially in countries such as Turkey and India.”

The report, which drew from Twitter’s own self-reported data, found that “a steep increase in the portion of requests that Twitter complies with in full.”

The data, Rest of World's **Russell Brandom** wrote, also does "not show a single request in which the company refused to comply, as it had done several times before the Musk takeover."

Wikipedia co-founder **Jimmy Wales**, whose nonprofit encyclopedia fought back against a since-rescinded ban on the site in Turkey, called out Musk for not resisting the country's orders.

"This is what it means to treat freedom of expression as a principle rather than a slogan," he tweeted.

While it's not uncommon for platforms to comply with court orders demanding content be taken down globally, New York Times tech reporter Ryan Mac suggested the move could create "a blueprint for repressive governments everywhere."

"If Twitter doesn't censor the content you want, simply threaten to cut off the service. Its owner just put it in writing," he tweeted.

As free speech advocates and members of the press have long pointed out, Musk's pledge to uphold "free speech" has always run the risk of conflicting with his comments about obeying local laws globally, given that many countries have moved to crack down on online speech.

And strictly adhering to local laws, as my colleagues **Naomi Nix** and **Gerry Shih** reported last year, could lead to greater safety risks "in countries where the government and powerful people frequently push social media giants to eliminate content" they don't like.

TEXT 12 - French regulator warns Twitter of legal duty to moderate misinformation, hate

The letter sent by Arcom gives Twitter a deadline of Nov. 24 by which it should confirm that it can still adhere to its obligations under French law to moderate content

By Annabelle Timsit, *The Washington Post*, November 22, 2022

France's digital regulator has asked Twitter to confirm that it can still meet its legal obligations to moderate harmful content and misinformation as the company undergoes a major reorganization, including layoffs of half the company's workforce, under the ownership of Elon Musk.

The head of Arcom, the French regulator for digital and audiovisual communication, expressed "deep concern" in a letter to Twitter about how the chaos and staff erosion that have characterized Musk's first weeks as owner and CEO since his \$44 billion purchase of Twitter last month could affect the company's ability to "maintain a safe environment for users of its service" in France.

Roch-Olivier Maistre, president of Arcom, noted in the letter, dated Friday, that Twitter "is one of the most used online platforms in France," and said the turmoil at the company "raises systemic issues regarding the faithfulness of democratic debate and the public safety."

He gave Twitter until Thursday to respond to Arcom's concerns.

Twitter did not respond to a request for comment from The Washington Post about the letter.

It comes as authorities in the United States and Europe vow to take a closer look at what is happening at Twitter, amid reports that a skeleton staff is operating the platform, and as Musk reinstates controversial accounts that were known for spreading harmful content — including those of former president Donald Trump and the Canadian psychologist Jordan Peterson — but whose suspensions led to criticism that Twitter was repressing its users' freedom of speech.

Twitter, the letter said, is responsible for fighting against "the manipulation of information" and "the dissemination of online hate speech" on its platform under French law, and Arcom is responsible for ensuring that Twitter adheres to its obligations, including on transparency. At the same time, Twitter should respect its users' freedom of expression, it said. If Twitter fails to implement measures to moderate online hate, the regulator has the power to put the platform on notice and, in some cases, impose fines of up to \$20.5 million, or 6 percent of its global revenue for the previous fiscal year. After Musk completed his takeover of Twitter in late October, he appointed himself CEO, fired top executives and cut half of the company's workforce. He said he would form a "content moderation council with widely diverse viewpoints" that would make future decisions on content moderation and suspended accounts.

The situation has affected Twitter's European operations: Over the weekend, the head of Twitter for France, Damien Viel, announced that he was leaving the company. He did not clarify whether he quit or was laid off.

"It's over," Viel tweeted.

The letter from Arcom highlighted reports of the layoffs, and questioned whether Twitter could meet its obligations under French and E.U. law with a reduced workforce, asking for clarification on "the short-term evolution of the human and technological resources devoted to compliance with these obligations."

Beyond the impact of the recent layoffs, the regulator also said it wants to ensure that Twitter can fully implement the Digital Services Act, a sweeping piece of legislation from the European Union that came into force last week and imposes transparency restrictions on tech companies.

Fueling speculation about the European response to Musk's takeover of Twitter was — appropriately enough — a tweet, after the purchase was finalized, from Thierry Breton, European commissioner for the internal market, in which Breton said Twitter would have to adhere to European rules.

In an editorial for the New York Times published last week, Yoel Roth, the former head of trust and safety at Twitter, wrote that "Twitter remains bound by the laws and regulations of the countries in which it operates," which could complicate Musk's professed wish for free speech to drive more decisions at Twitter. "Regulators have significant tools at their disposal to enforce their will on Twitter and on Mr. Musk," Roth added, citing the E.U.'s Digital Services Act. The departure of Roth and other employees tasked with security and compliance at Twitter after Musk's takeover prompted the Federal Trade Commission in the United States to warn that it was prepared to step in to verify that the company was respecting the terms of an agreement to ensure data security and privacy for users of the platform.

As The Post reported, Musk told Twitter employees in an email that the company "will do whatever it takes to adhere to both the letter and spirit" of its agreement with the FTC. "The same goes for any other government regulatory matters where Twitter operates," he added.

TEXT 13 - Désinformation en ligne : Twitter quitte le code de l'Union européenne, annonce Thierry Breton

Par Emmanuel Egloff, *Le Figaro*, 27/05/2023

Malgré le retrait du réseau social de ce code de bonne conduite, «ses obligations demeurent», a tenu à rappeler le commissaire européen à l'Industrie.

L'adhésion au code de bonnes pratiques de l'Union européenne contre la désinformation en ligne (see **HERE**) est volontaire. Créé en 2018, il comprend 128 mesures visant à mieux coopérer avec les fact-checkeurs et à priver de publicité les sites diffusant de fausses informations. Et toutes les grandes plateformes (Meta - propriétaire de Facebook -, Google, Microsoft, TikTok...) y adhèrent. Mais plus Twitter. Le réseau social, racheté en octobre dernier par Elon Musk, a décidé de se désengager de ce code de bonnes pratiques. Il est vrai que, depuis ce rachat, les règles de modération de la plateforme ont été largement assouplies.

La situation ne va toutefois pas rester longtemps en l'état. À partir du 25 août, le DSA, règlement européen sur le numérique, s'appliquera. Il sera obligatoire et va contraindre toutes les plateformes à lutter contre la désinformation. «*Nos équipes sont prêtes à faire appliquer la loi* » a déclaré, lundi 29 mai, Thierry Breton, le commissaire européen en charge du Marché intérieur.

Et les sanctions sont dissuasives. Le DSA prévoit des amendes pouvant aller jusqu'à 6% de leur chiffre d'affaires mondial pour les entreprises qui ne respecteraient pas leurs obligations. En tout dernier recours, «*il sera possible de demander au tribunal une suspension temporaire de leur service* », précise la Commission européenne. Jean-Noël Barrot, le ministre délégué français chargé de la Transition numérique, a enfoncé le clou. «*À partir du 25 août prochain, de nouvelles règles s'appliqueront, a-t-il expliqué. Si Twitter ne s'y conforme pas, il s'exposera à des sanctions pouvant aller jusqu'à son blocage dans l'Union européenne.* » Une déclaration que le ministre a choisi de faire... sur Twitter.

TEXT 14 - Linda Yaccarino, une nouvelle PDG de Twitter déjà controversée

Elon Musk a confirmé avoir nommé sa remplaçante à la tête du réseau social, la directrice de la publicité et des partenariats de NBCUniversal, qui a notamment conseillé l'administration Trump.

Elon Musk avait bien pris garde de ménager le suspense en annonçant, jeudi 11 mai, avoir « *recruté une nouvelle PDG pour X/Twitter* », qui devrait prendre ses fonctions dans six semaines, mais sans préciser son identité. Quelques heures plus tard pourtant, *The Wall Street Journal* puis *The Washington Post* révélaient le nom de la nouvelle patronne du réseau social, et vendredi 12 mai, Elon Musk mettait fin au suspense en confirmant ce nom : Linda Yaccarino sera la nouvelle directrice générale de Twitter. « *Elle va se concentrer principalement sur les affaires, tandis que je vais m'occuper du design du produit et des nouvelles technologies* », a-t-il précisé dans son tweet.

M^{me} Yaccarino était jusqu'ici directrice de la publicité et des partenariats du conglomerat de médias américains NBCUniversal, filiale de Comcast, qui regroupe les chaînes NBC, plusieurs studios de cinémas et des services de streaming. Son départ a été officialisé par le groupe dans un communiqué publié vendredi.

Linda Yaccarino a fait toute sa carrière dans les médias, d'abord au sein du groupe Turner, puis chez NBCUniversal, où elle occupe des postes de direction de la publicité depuis dix ans. Sans expérience directe dans la Silicon Valley, elle connaît cependant très bien le monde des réseaux sociaux, notamment parce qu'elle a construit au fil des années de multiples partenariats avec ces entreprises, dont Twitter. A la mi-avril, M^{me} Yaccarino avait reçu Elon Musk pour une interview dans une conférence sur le marketing. Selon les informations du site *Axios*, ils se sont revus le 9 mai en marge d'une autre conférence, deux jours avant qu'Elon Musk annonce avoir trouvé sa « *remplaçante* ».

Après avoir lancé un sondage demandant aux utilisateurs s'il devait laisser la main – ces derniers lui avaient alors majoritairement répondu « oui » –, l'homme d'affaires avait annoncé, en décembre, s'être mis à la recherche d'un nouveau PDG pour Twitter. Le temps passé par Elon Musk à la tête du réseau social, et son comportement souvent erratique à ce poste, ont indirectement entraîné d'importantes chutes des actions Tesla, les investisseurs craignant qu'il délaisse son entreprise la plus rentable.

Marge de manœuvre limitée

Reste à savoir quelle sera la marge de manœuvre réelle de M^{me} Yaccarino. Elon Musk a, en effet, également annoncé qu'il resterait CTO (*chief technical officer*, le poste de direction qui encadre l'ensemble des équipes techniques) de Twitter ainsi que président du conseil d'administration. En pratique, il conservera donc un pouvoir très important sur toutes les orientations de la plate-forme.

Le choix d'une spécialiste de la publicité en ligne est par ailleurs en partie en contradiction avec les dernières décisions de M. Musk, qui semblait vouloir miser l'avenir du réseau social sur le service payant Twitter Blue davantage que sur la publicité. Depuis sa prise de contrôle du réseau social, les recettes publicitaires se sont effondrées, notamment en raison d'une fuite des grandes sociétés parmi ses annonceurs, effrayées par ses décisions limitant la modération sur la plate-forme ainsi que le licenciement de l'essentiel des modérateurs du site comme de très nombreux commerciaux chargés de vendre la publicité sur Twitter.

Le choix de Linda Yaccarino a également soulevé des questions plus politiques. Plutôt discrète sur ses convictions personnelles, la nouvelle directrice du réseau social avait cependant été choisie par l'administration Trump pour diriger un conseil présidentiel sur le sport et la nutrition. Plusieurs observateurs et médias ont par ailleurs noté qu'elle suit l'activité sur Twitter de très nombreux comptes de la droite dure américaine, pour certains appréciés d'Elon Musk. Dans le même temps, l'annonce de sa nomination a provoqué de vives réactions de l'extrême droite complotiste américaine, notamment en raison de son soutien à la vaccination contre le Covid-19, mais aussi parce que M^{me} Yaccarino dirige un groupe de travail au sein du Forum économique mondial, objet de multiples théories du complot.

The Long Read - Social media's struggle with self-censorship

Tech giants are removing more content, but are they making the right choices?

The Economist, Oct 22nd 2020

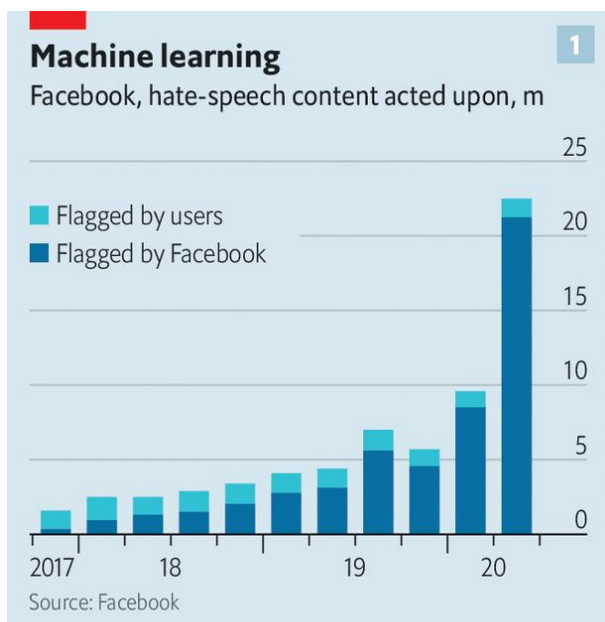
Within hours of the publication of a *New York Post* article on October 14th, Twitter users began receiving strange messages. If they tried to share the story—a dubious “exposé” of emails supposedly from the laptop of Hunter Biden, son of the Democratic presidential nominee—they were told that their tweet could not be sent, as the link had been identified as harmful. Many Facebook users were not seeing the story at all: the

If the companies had hoped that by burying or blocking the story they would stop people from reading it, the bet did not pay off. The article ended up being the most-discussed story of the week on both platforms—and the second-most talked-about story was the fact that the social networks had tried to block it. The *Post* called it an act of modern totalitarianism, carried out “not [by] men in darkened cells driving screws under the fingernails of dissidents, but Silicon Valley dweebs.” Republican senators vowed to extract testimony on anticonservative bias from Mark Zuckerberg and Jack Dorsey, the dweebs-in-chief of, respectively, Facebook and Twitter.

The tale sums up the problem that social networks are encountering wherever they operate. They set out to be neutral platforms, letting users provide the content and keeping their hands off editorial decisions. Twitter executives used to joke that they were “the free-speech wing of the free-speech party”. Yet as they have become more active at algorithmically ranking the content that users upload, and moderating the undesirable stuff, they

have edged towards being something more like publishers. Mr Zuckerberg says he does not want to be an “arbiter of truth”. The *Post* episode fed the suspicion of many that, willingly or not, that is precisely what he is becoming.

America's fractious election campaign has only made more urgent the need to answer the unresolved questions about free expression online. What speech should be allowed? And who should decide? Rasmus Nielsen of the Reuters Institute at Oxford University describes this as a “constitutional moment” for how to regulate the private infrastructure that has come to support free expression around the world.



The Economist

Social networks have been on the mother of all clean-ups. Facebook's removal of hate speech has risen tenfold in two years (see chart 1). It disables some 17m fake accounts every single day, more than twice the number three years ago. YouTube, a video platform owned by Google with about 2bn monthly users, removed 11.4m videos in the past quarter, along with 2.1bn user comments, up from just 166m comments in the second quarter of 2018. Twitter, with a smaller base of about 350m users, removed 2.9m tweets in the second half of last year, more than double the amount a year earlier. TikTok, a Chinese short-video upstart, removed 105m clips in the first half of this year, twice as many as in the previous six months (a jump partly explained by the firm's growth).

Artificial intelligence has helped to make such a clean-up possible. Most offending content is taken down before any user has had a chance to flag it. Some lends itself readily to policing with machines: more than 99% of the child-nudity posts Facebook takes down are removed before anyone has reported them, but most of the bullying or harassment is flagged by users rather than robots. Two years ago Facebook's ai removed a post referring to "merciless Indian Savages", before human moderators realised it was a quote from the Declaration of Independence. Facebook now employs about 15,000 people to moderate content. In May the company agreed to pay \$52m to 11,250 moderators who developed post-traumatic stress disorder from looking at the worst of the internet.

Discussions about free speech that may once have seemed abstract have become all too practical—the murder of Samuel Paty near Paris last week being the latest shocking reminder. Social networks tightened

their policies on terrorism after Islamist attacks in Europe in 2015 and an anti-Muslim rampage in New Zealand last year, which was live-streamed on Facebook and shared on YouTube. The American election and Brexit referendum of 2016 forced them to think again about political communication. Twitter banned all political ads last year, and Facebook and Google have said they will ban them around the time of this year's election on November 3rd.

The companies have also improved their scrutiny of far-flung countries, after criticism of their earlier negligence in places such as Myanmar, where Facebook played a "determining role" in the violence against Rohingya Muslims, according to the un (see [article](#)). This week Facebook announced that it had hired more content-reviewers fluent in Swahili, Amharic, Zulu, Somali, Oromo and Hausa, ahead of African elections. Its ai is learning new languages, and hoovering up rule-breaking content as it does so.

The room where it happens

Some tech bosses have been rethinking their approach to the trade-offs between free expression and safety. Last October, in a speech at Georgetown University, Mr Zuckerberg made a full-throated defence of free speech, warning: "More people across the spectrum believe that achieving the political outcomes they think matter is more important than every person having a voice. I think that's dangerous." Yet this year, as misinformation about covid-19 flourished, Facebook took a harder line on fake news about health, including banning anti-vaccination ads. And this month it banned both Holocaust denial and groups promoting QAnon, a crackpot conspiracy.

The pressure from the media is to "remove more, remove more, remove more", says one senior tech executive. But in some quarters unease is growing that the firms are removing too much. In America this criticism comes mostly from the right, which sees Silicon Valley as a nest of liberals. It is one thing to zap content from racists and Russian trolls; it is another to block the *New York Post*, one of America's highest-circulation newspapers, founded by Alexander Hamilton (who admittedly might not have approved of its current incarnation, under Rupert Murdoch).

Elsewhere, liberals worry that whistle-blowing content is being wrongly taken down. YouTube removed footage from users in Syria that it deemed to break its guidelines on violence, but which was also potential evidence of war crimes. Until last year TikTok's guidelines banned criticism of systems of government

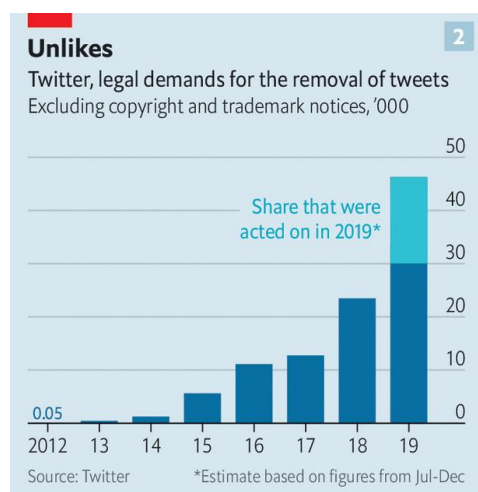
and “distortion” of historical events including the massacre near Tiananmen Square.

Where both camps agree is in their unease that it is falling to social networks to decide what speech is acceptable. As private companies they can set their own rules about what to publish (within the confines of the laws of countries where they operate). But they have come to play a big role in public life. Mr Zuckerberg himself compares Facebook to a “town square”.

Rival social networks promising truly free speech have struggled to overcome the network effects enjoyed by the incumbents. One, Gab, attracted neo-Nazis. Another, Parler, has been promoted by some Republican politicians but so far failed to take off. (It is also grappling with free-speech dilemmas of its own, reluctantly laying down rules including no sending of photos of fecal matter.) Outside China, where Facebook does not operate, four out of ten people worldwide use the platform; WhatsApp and Instagram, which it also owns, have another 3bn or so accounts between them. “Frankly, I don’t think we should be making so many important decisions about speech on our own either,” Mr Zuckerberg said in his Georgetown speech.

Say no to this

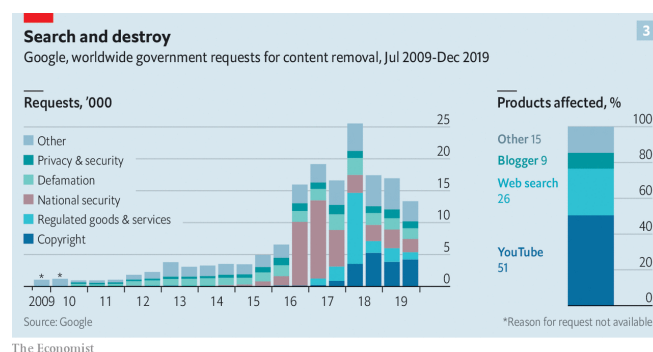
Bill Clinton once said that attempting to regulate the internet, with its millions of different sites, would be “like trying to nail Jell-O to the wall”. But the concentration of the social-media market around a few companies has made the job easier.



The Economist

Twitter has faced steep growth in the number of legal requests for content removal, from individuals as well as governments (see chart 2). Last year Google received 30,000 requests from governments to remove pieces of content, up from a couple of thousand requests ten years

ago (see chart 3). And Facebook took down 33,600 pieces of content in response to legal requests. They included a Photoshopped picture of President Emmanuel Macron in pink underwear, which French police wanted removed because it broke a law from 1881 restricting press freedom.



In America the government is prevented from meddling too much with online speech by the First Amendment. Section 230 of the Communications Decency Act gives online platforms further protection, exempting them from liability for the content they publish. But carve-outs to this exemption are growing. Firms cannot avoid responsibility for copyright infringements, posts that break federal criminal law, or which enable sex trafficking. The latter exemption, made in 2018, had an impact on speech that was greater than its drafting implied: sites including Tumblr and Craigslist concluded that, rather than risk prosecution, they would stop publishing adult material of all sorts.

In Europe regulation has gone further. In 2014 the European Court of Justice (ecj) established the “right to be forgotten” when it found in favour of a Spanish man who wanted Google to remove old references to his history of indebtedness. Since then Google has fielded requests for about half a million urls to be removed each year, and granted about half of them. Last year the ecj ruled that European countries could order Facebook to remove content worldwide, not just for users within their borders. The European Audiovisual Media Services Directive requires online video services to take “appropriate measures” to protect viewers from harmful or illegal content, including setting up age checks. The European Commission is to publish a Digital Services Act, expected to impose further obligations on internet companies.

National governments have also set their own rules, notably Germany, whose Network Enforcement Act of 2017 threatens platforms with fines of up to €50m (\$60m) if they fail to take down illegal content within 24 hours of notification. In response Facebook opened a new moderation centre in Germany. The trouble with

privatising the enforcement of the law in this way, points out Mr Nielsen, is that the companies have big incentives to err on the side of caution. A judge may use discretion to ignore rules on speech that are seldom applied (such as a German law that until recently banned insulting a foreign head of state). But a social-media company has no reason to risk ignoring a law.

Who tells your story

Some governments are leaning on social networks to remove content that may be legal. The social-media platforms have their own rules that go further than most governments'. A ban on material that could interfere with "civic integrity" may sound like something from communist China; it is actually in Twitter's rules. London's Metropolitan Police has a unit that scours platforms for terrorism-related content, which it "requests" be taken down for breaching the platform's terms of service—even though the material may not break any law.

"Authoritarian governments are taking cues from the loose regulatory talk among democracies," writes David Kaye, a former UN special rapporteur on free expression. Last year Singapore passed what it described as an anti-fake-news law, banning the online publication of lies that could harm the public interest. Thailand has enforced its *lèse-majesté* laws online, in August ordering Facebook to block a critical group called Royalist Marketplace, which with more than 1m members was one of the largest on the platform. (Facebook complied, but is suing the Thai government for breaking human-rights law.)

If neither governments nor executives make reliable custodians of free speech, what can be done to keep the internet a tolerable place while protecting freedom of expression? An increasingly common answer in Silicon Valley is to draw a distinction between freedom of speech and "freedom of reach": leave posts up, but make them less visible and viral.

Last year YouTube changed its algorithm so that videos that were borderline cases for deletion were recommended less often. After the bombings of churches and hotels in Sri Lanka at Easter in 2019, Facebook prevented the resharing of posts by friends of friends, to stop inflammatory content travelling too far or fast; this rule is in place in Ethiopia and Myanmar. Twitter has tried to stop people from mindlessly sharing fake news by prompting them to read articles before they retweet them. Platforms are adding more labels to content, warning users that it is misleading.

Another idea gaining momentum is that firms should make their data available for audit just as listed companies must open up their accounts. Their internal processes could also be more transparent. At Facebook there is an odd tension between its earnest approach to policymaking, with fortnightly "mini-legislative sessions", and the fact that every month Mr Zuckerberg personally takes a handful of the hardest decisions on content moderation. Treating the big calls as "corner-office decisions" is a mistake, believes Mr Kaye: better for companies to say, "We have these rules, we're going to apply them neutrally. And we don't want that process to be corrupted by political pressure."

Facebook took a step towards such a system on October 22nd with the launch of its Oversight Board, a watchdog made up of 20 members of the great and good who will scrutinise its moderation decisions and issue binding rulings. The board's scope is narrower than some had hoped. It can consider only whether deleted posts should be reinstated. It merely applies Facebook's rules, rather than setting them. It cannot consider posts that have been algorithmically demoted, as opposed to deleted. So some of the most prominent recent controversies—Facebook's decision to leave up a contentious post by Donald Trump, its removal of QAnon, its reversal on Holocaust denial and its demotion of the *Post* story—are outside the board's jurisdiction.

History has its eyes on you

Yet as Alan Rusbridger, a former *Guardian* editor and member of the new board, puts it, it is a "revolutionary thought". "A company that has notoriously been very reluctant to surrender control on anything has handed over...the power to make some pretty consequential decisions on its behalf," he says. He hopes the board will get more powers over time. Facebook says this is premature. But Sir Nick Clegg, its head of global affairs, hopes the board's remit might one day expand to consider cases submitted by other social networks.

Others have similar ideas. Article 19, a free-speech lobby group, has suggested that platforms could outsource their moderation decisions to non-governmental "social-media councils", something like the press watchdogs that in many countries hold newspapers to a voluntary code.

For now, the social networks have to get through perhaps the hardest fortnight in their short history. They face the possibility of having to deploy content-moderation tools developed for fragile, emerging democracies in their home country. Facebook removed

120,000 pieces of content aimed at voter suppression in America in the past quarter. The *New York Post* affair does not bode well for how the companies might handle the fallout from a contested election. “When they appeared to depart from their policies they opened

themselves up to the very charges of bias that followed,” says Evelyn Douek of Harvard Law School. As the election approaches, they need to “tie themselves to a mast” of clear rules, she says. A storm is coming.

Elon Musk Is Right That Twitter Should Follow the First Amendment

A long history of free-speech jurisprudence backs him up.

By [Jeffrey Rosen](#)



Getty; The Atlantic

The Atlantic, MAY 2, 2022

Elon Musk, in his effort to buy Twitter, signaled that under his ownership, the company would allow all speech that the First Amendment protects. “By ‘free speech,’ I simply mean that which matches the law,” he tweeted on April 26. “I am against censorship that goes far beyond the law.”

Many commentators were quick to point out that, as a private company, Twitter is not required to follow the First Amendment, which applies only to federal and state governments. And Musk has further been criticized by those who fear that harmonizing Twitter’s content rules with First Amendment doctrine would lead to an explosion of hate speech, misinformation, and incendiary statements, content that Twitter currently moderates.

This deregulatory approach would make Twitter an outlier among the social-media companies; at the moment, Twitter, like Facebook and Google, has chosen to adopt content rules stricter than First Amendment standards. Facebook, for example, prohibits “hate speech” or “attacks,” which it defines as “violent or dehumanizing speech, harmful stereotypes, statements of inferiority, expressions of contempt, disgust or dismissal, cursing and calls for exclusion or segregation.” By contrast, the First Amendment generally protects hate speech unless it is intended, and likely, to cause imminent injury. Twitter’s current definition of “hateful conduct,” although narrower than Facebook’s, still falls short of First Amendment standards.

But Musk’s position is, in fact, convincing. Although private companies are not required to follow the First Amendment, nothing prevents them from doing so voluntarily. And in Twitter’s case in particular, there are strong reasons to believe that the First Amendment should presumptively govern. All four of the main principles that

have historically guided the Supreme Court in interpreting the First Amendment apply just as powerfully to social-media platforms as they do to governments.

What are these First Amendment first principles? Justice Louis Brandeis expressed all four in his opinion in *Whitney v. California*, a 1927 case that involved a woman convicted of making a speech at a Communist Party meeting in support of anti-lynching laws. Here is Brandeis's crucial paragraph, in which he drew heavily upon Thomas Jefferson's "Bill for Establishing Religious Freedom":

Those who won our independence believed that the final end of the state was to make men free to develop their faculties, and that in its government the deliberative forces should prevail over the arbitrary. They valued liberty both as an end and as a means. They believed liberty to be the secret of happiness and courage to be the secret of liberty. They believed that freedom to think as you will and to speak as you think are means indispensable to the discovery and spread of political truth; that without free speech and assembly discussion would be futile; that with them, discussion affords ordinarily adequate protection against the dissemination of noxious doctrine; that the greatest menace to freedom is an inert people; that public discussion is a political duty; and that this should be a fundamental principle of the American government.

Based on these four principles—freedom of conscience, democratic accountability, discovery of truth, and democratic self-government—Brandeis articulated the First Amendment test that the Supreme Court would later adopt in 1969: The government can regulate speech under the First Amendment only when the speech is intended, and likely, to cause imminent and serious injury. Intent, imminence, and severity are crucial components of this test. (Twitter currently forbids accounts whose “primary purpose is inciting harm toward others” and could continue to prohibit and remove imminent threats, targeted harassment, defamation, and other speech that can be defined as illegal under a rigorous First Amendment standard.)

As Brandeis explained in *Whitney*:

Fear of serious injury cannot alone justify suppression of free speech and assembly. Men feared witches and burnt women. It is the function of speech to free men from the bondage of irrational fears. To justify suppression of free speech there must be reasonable ground to fear that serious evil will result if free speech is practiced. There must be reasonable ground to believe that the danger apprehended is imminent.

Brandeis's test was based on his Jeffersonian faith in the power of what he called “free and fearless reasoning” to expose falsehood through public discussion. “If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education,” Brandeis wrote, “the remedy to be applied is more speech, not enforced silence. Only an emergency can justify repression.” As a result, the United States Supreme Court now protects free speech more vigorously than any other judiciary in the world.

Brandeis's faith in reason—and his four justifications for protecting all speech not intended to and likely to cause violence—is being questioned in our social-media age. Twitter, Facebook, and other platforms have accelerated public discourse to warp speed, creating virtual versions of the mob. These companies are based on a business model that's now being called “enrage to engage.” Inflammatory posts based on passion travel further and faster than arguments based on reason. Because of this, some critics are suggesting that America's free-speech tradition should be reconsidered or amended.

These arguments are unpersuasive, and Musk is correct to reject them. In fact, all four of Brandeis's arguments weigh heavily against any central authority, whether government or Twitter, exercising the power to monitor free speech. Let's consider each in turn.

1. Freedom of conscience is an unalienable right because people can think only for themselves.

If freedom means anything, it is the freedom to “develop our faculties,” Brandeis wrote in *Whitney*, by which he meant our faculties of reason and deliberation. Brandeis and Jefferson rightly believed that freedom of conscience is, by definition, an unalienable right, one that can't be surrendered or alienated to government or private actors,

because our opinions, as Jefferson wrote in the draft of his Virginia Bill for Establishing Religious Freedom, “follow involuntarily the evidence proposed to [our] minds.” We can’t give anyone—presidents, priests, teachers, or social-media magnates and their content monitors—the power to tell us what to think, even if we wanted to, because we are endowed as human beings with the capacity to reason and therefore can’t help thinking for ourselves. And a crucial element of thinking for ourselves is deciding for ourselves what evidence is proper for us to consider. Once a public or private regulator gets in the business of deciding which opinions are true or false, or what kind of information is good for people to access, that regulator infringes on the right of all individuals to form opinions with what Jefferson called “the illimitable freedom of the human mind.”

2. Free speech makes representatives accountable to We the People.

As Brandeis wrote, in a democracy “the deliberative forces should prevail over the arbitrary.” This requires people to be able to speak truth to power. The right of the people to criticize all political figures, and the right of political figures to communicate with the people, is crucial to a functioning democracy. Jefferson and James Madison attacked the Alien and Sedition Acts of 1798, which criminalized criticism of the Federalist president John Adams, but not of the Republican vice president Thomas Jefferson, on the grounds that favoring some politicians over others infringed on people’s right “of freely examining public characters and measures, and of free communication among the people thereon, which has ever been justly deemed, the only effectual guardian of every other right.” For this reason, Musk is correct to question lifetime bans of politicians such as Donald Trump, which are administered in a way that discriminates based on viewpoint, which First Amendment doctrine explicitly forbids. Trump has been banned permanently from Twitter while other leaders who have tweeted similarly inflammatory statements that arguably fall short of the legal standard for incitement have not been deplatformed, such as Ayatollah Khamenei’s tweets calling for the eradication of Israel. The possibility of viewpoint discrimination inherent in giving social-media platforms the power to ban some public officials but not others means that Musk is right to embrace the position that the new Facebook free-speech advisory board has reached as well: Temporary time-outs are acceptable, but lifetime bans are not.

3. Free speech is the best way to ensure the “discovery and spread of political truth.”

Brandeis and Jefferson had faith in the power of reasoned deliberation to distinguish truth from error over time. As Jefferson put it, “We are not afraid to follow truth wherever it may lead, nor to tolerate any error so long as reason is left free to combat it.” At the moment, Twitter empowers its content monitors to flag “misinformation,” defined as tweets that lack “context.” But this vague standard has been applied to take down posts that later turned out to be true. Referring to Twitter’s decision to suspend the *New York Post* for tweeting articles about Hunter Biden that eventually proved to be accurate, Musk rightly tweeted that “suspending the Twitter account of a major news organization for publishing a truthful story was obviously incredibly inappropriate.”

Ever since the English philosopher John Milton, in *Areopagitica*, denounced the censorship of books as inappropriate because identifying censors infallible enough to consistently and reliably distinguish truth from error would be impossible, free-speech defenders have recognized that truth must ultimately emerge bottom-up from public discussion among the people, rather than top-down from a paternalistic (and possibly self-interested) regulator. Free citizens in a liberal democracy shouldn’t trust any centralized authority—public or private—to make decisions about what books, music, and other content they can safely be exposed to.

4. Free speech allows the public discussion necessary for democratic self-government.

Jefferson and Brandeis believed that in a democracy, all citizens have an equal right and responsibility to exercise their freedom of conscience. When social-media platforms presume to decide in advance what sorts of political candidates or information are safe for the people to evaluate, they deny the people their right to make that decision on their own. In the process, they weaken the public’s ability to make the political choices that are the foundation of a functioning democracy.

Elon Musk’s position that Twitter should abide by the First Amendment is a radical one—at the moment, no other major social-media platform, with arguably the exception of Reddit, has chosen voluntarily to embrace First Amendment standards. But for the reasons Jefferson and Brandeis recognized, Musk is entirely correct. It doesn’t matter whether the president or a prominent social-media company is presuming to tell us what to think or restricting the information we receive in the interest of protecting us. In the end, all of us have a right and a responsibility to “think as [we] will and to speak as [we] think,” as Brandeis put it. In other words, we have no choice but to think for ourselves.

[Jeffrey Rosen](#) is a contributing writer for *The Atlantic*, President & CEO of the National Constitution Center, and a law professor at George Washington University. He is the author of [*Conversations With RBG: Justice Ruth Bader Ginsburg on Life, Love, Liberty, and Law*](#).